# *Inequality with ordinal data*

Martyna Kobus

Polish Academy of Sciences, Institute of Economics

January 14, 2016

# Plan

- Motivation
- Univariate distributions: measurement of health inequality, educational inequality, happiness inequality etc.
- Bivariate distributions: socioeconomic inequalities in health (education-health gradient, income-health gradient), mobility measurement, more generally measuring interdependence
- Summary and future research

# *Plan*

- Motivation
- Univariate distributions: measurement of health inequality, educational inequality, happiness inequality etc.
- Bivariate distributions: socioeconomic inequalities in health (education-health gradient, income-health gradient), mobility measurement, more generally measuring interdependence
- Summary and future research

# Plan

- Motivation
- Univariate distributions: measurement of health inequality, educational inequality, happiness inequality etc.
- Bivariate distributions: socioeconomic inequalities in health (education-health gradient, income-health gradient), mobility measurement, more generally measuring interdependence
- Summary and future research

# Plan

- Motivation
- Univariate distributions: measurement of health inequality, educational inequality, happiness inequality etc.
- Bivariate distributions: socioeconomic inequalities in health (education-health gradient, income-health gradient), mobility measurement, more generally measuring interdependence
- Summary and future research

## *Motivation*

- There has been a growing interest in the last decade in the treatment of ordinal data when measuring inequality, poverty or welfare. This comes from both literature and policymaking.
- Standard literature on inequality measurement focuses on the inequality of attributes that are measurable and comparable among different units of measurement (individuals, groups, regions, countries etc.) e.g. income.

# *Motivation*

- There has been a growing interest in the last decade in the treatment of ordinal data when measuring inequality, poverty or welfare. This comes from both literature and policymaking.
- Standard literature on inequality measurement focuses on the inequality of attributes that are measurable and comparable among different units of measurement (individuals, groups, regions, countries etc.) e.g. income.

- Yet in the last decades researchers have acknowledged that, while important, income cannot be treated as sole indicator of well-being (Atkinson and Bourguignon 1982; Sen 1973, 1987; Maasoumi 1986; Tsui 1995; Gajdos and Weymark 2005; Duclos, Sahn and Younger 2011). Well-being is a multidimensional concept (Stiglitz, Sen and Fitoussi 2008).
- Empirical analyses increasingly include non-income dimensions of well-being e.g. health.

- Yet in the last decades researchers have acknowledged that, while important, income cannot be treated as sole indicator of well-being (Atkinson and Bourguignon 1982; Sen 1973, 1987; Maasoumi 1986; Tsui 1995; Gajdos and Weymark 2005; Duclos, Sahn and Younger 2011). Well-being is a multidimensional concept (Stiglitz, Sen and Fitoussi 2008).
- Empirical analyses increasingly include non-income dimensions of well-being e.g. health.

# Motivation

- Some governments (e.g. UK, France, Canada, Japan) and international organizations have been responding to economists' urges to incorporate a multidimensional perspective on well-being (progress, quality of life), inequality, and poverty.
- In 2010 British Prime Minister 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. Cameron described monitoring people's well-being as one of central political issues of our time.
- The UK Office for National Statistics has a program "Measuring what matters" in which they are developing new measures of national well-being.
- In 2008 French government created a renowned Stiglitz-Sen-Fitousi commission in order to better tackle the problem of measuring social and economic performance.

# Motivation

- Some governments (e.g. UK, France, Canada, Japan) and international organizations have been responding to economists' urges to incorporate a multidimensional perspective on well-being (progress, quality of life), inequality, and poverty.
- In 2010 British Prime Minister 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. Cameron described monitoring people's well-being as one of central political issues of our time.
- The UK Office for National Statistics has a program "Measuring what matters" in which they are developing new measures of national well-being.
- In 2008 French government created a renowned Stiglitz-Sen-Fitousi commission in order to better tackle the problem of measuring social and economic performance.

# *Motivation*

- Some governments (e.g. UK, France, Canada, Japan) and international organizations have been responding to economists' urges to incorporate a multidimensional perspective on well-being (progress, quality of life), inequality, and poverty.
- In 2010 British Prime Minister 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. Cameron described monitoring people's well-being as one of central political issues of our time.
- The UK Office for National Statistics has a program "Measuring what matters" in which they are developing new measures of national well-being.
- In 2008 French government created a renowned Stiglitz-Sen-Fitousi commission in order to better tackle the problem of measuring social and economic performance.

# *Motivation*

- Some governments (e.g. UK, France, Canada, Japan) and international organizations have been responding to economists' urges to incorporate a multidimensional perspective on well-being (progress, quality of life), inequality, and poverty.
- In 2010 British Prime Minister 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. Cameron described monitoring people's well-being as one of central political issues of our time.
- The UK Office for National Statistics has a program "Measuring what matters" in which they are developing new measures of national well-being.
- In 2008 French government created a renowned Stiglitz-Sen-Fitousi commission in order to better tackle the problem of measuring social and economic performance.

# Motivation

- In May 2011, in its 50th Anniversary Week, 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.
- Some NGOs become involved too e.g. Social Progress Imperative develops Social Progress Index and aims to "solve the world's most pressing challenges by redefining how the world measures success".

# *Motivation*

- In May 2011, in its 50th Anniversary Week, 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.
- Some NGOs become involved too e.g. Social Progress Imperative develops Social Progress Index and aims to “solve the world’s most pressing challenges by redefining how the world measures success".

# Motivation

- Many non-income wellbeing dimensions are available in surveys in the form of ordinal data. Such is the case with self-reported health status, educational attainment, life satisfaction, living conditions and many others. Mostly in health economics, happiness economics, educational economics, and development economics.
- To be precise, we deal with data that are ordinal and discrete. Ordinal - invariance with respect to monotone transformations (vs. cardinality - particular numbers are meaningful). Discrete - variables are concentrated on a fixed number of points (vs. continuous variables which accord a particular value with probability zero).
- There are numerous measurement problems with ordinal data.

# Motivation

- Many non-income wellbeing dimensions are available in surveys in the form of ordinal data. Such is the case with self-reported health status, educational attainment, life satisfaction, living conditions and many others. Mostly in health economics, happiness economics, educational economics, and development economics.
- To be precise, we deal with data that are ordinal and discrete. Ordinal - invariance with respect to monotone transformations (vs. cardinality - particular numbers are meaningful). Discrete - variables are concentrated on a fixed number of points (vs. continuous variables which accord a particular value with probability zero).
- There are numerous measurement problems with ordinal data.

# Motivation

- Many non-income wellbeing dimensions are available in surveys in the form of ordinal data. Such is the case with self-reported health status, educational attainment, life satisfaction, living conditions and many others. Mostly in health economics, happiness economics, educational economics, and development economics.
- To be precise, we deal with data that are ordinal and discrete. Ordinal - invariance with respect to monotone transformations (vs. cardinality - particular numbers are meaningful). Discrete - variables are concentrated on a fixed number of points (vs. continuous variables which accord a particular value with probability zero).
- There are numerous measurement problems with ordinal data.

# Motivation

- There is another strand of literature which is not related to well-being measurement, but uses ordinal data a lot, namely so called socioeconomic inequalities in health (Wagstaff et al. 1991, Kakwani et al. 1997, O'Donnell et al. 2008, van Doorslaer and Jones 2003).
- This literature involves the study of the relationship between income/SES/education and health. We focus on pure measurement issues here (no causality etc.)
- In a sense, the study of univariate distributions measures total dispersion in health, irrespective of socioeconomic factors. To include socioeconomic dimensions (like income or education) one necessarily deals with (at least) bivariate distributions.

# Motivation

- There is another strand of literature which is not related to well-being measurement, but uses ordinal data a lot, namely so called socioeconomic inequalities in health (Wagstaff et al. 1991, Kakwani et al. 1997, O'Donnell et al. 2008, van Doorslaer and Jones 2003).
- This literature involves the study of the relationship between income/SES/education and health. We focus on pure measurement issues here (no causality etc.)
- In a sense, the study of univariate distributions measures total dispersion in health, irrespective of socioeconomic factors. To include socioeconomic dimensions (like income or education) one necessarily deals with (at least) bivariate distributions.

- There is another strand of literature which is not related to well-being measurement, but uses ordinal data a lot, namely so called socioeconomic inequalities in health (Wagstaff et al. 1991, Kakwani et al. 1997, O'Donnell et al. 2008, van Doorslaer and Jones 2003).
- This literature involves the study of the relationship between income/SES/education and health. We focus on pure measurement issues here (no causality etc.)
- In a sense, the study of univariate distributions measures total dispersion in health, irrespective of socioeconomic factors. To include socioeconomic dimensions (like income or education) one necessarily deals with (at least) bivariate distributions.

- The standard procedure of dealing with ordinal variables is to try to cardinalize them (O'Donell et al. 2008; Van Doorslaer and Jones 2003). The cardinalization may come from a simple assignment of a sequence of numbers (Deaton and Paxson 1998) but also via assuming a distribution of a latent variable or through tying a distribution of an ordinal variable to a distribution of a cardinal variable.
- Allison and Foster (2004) show that cardinalization is a flawed procedure. Variations in the scale that cardinalizes an ordinal variable may reverse the ranking of distributions based on means, variance, coefficient of variation (Lazar and Silber 2013) or standard inequality measures like Gini index (Kobus 2015).

# Univariate distributions: the measurement problem

- The standard procedure of dealing with ordinal variables is to try to cardinalize them (O'Donell et al. 2008; Van Doorslaer and Jones 2003). The cardinalization may come from a simple assignment of a sequence of numbers (Deaton and Paxson 1998) but also via assuming a distribution of a latent variable or through tying a distribution of an ordinal variable to a distribution of a cardinal variable.
- Allison and Foster (2004) show that cardinalization is a flawed procedure. Variations in the scale that cardinalizes an ordinal variable may reverse the ranking of distributions based on means, variance, coefficient of variation (Lazar and Silber 2013) or standard inequality measures like Gini index (Kobus 2015).

- More recently, Bond and Lang (2014) arrive at the same conclusion in a slightly different way.
- If the latent distribution can be represented by a a two-parameter distribution (e.g. cdf being a function of mean and variance) with different parameters for each group being compared - like in any estimation using ordered probit or ordered logit - then the CDFs of two groups will always cross except for a zero-probability event where the variances are the same, so there cannot be FOD which is necessary for the mean of two distributions to be invariant to monotonic transformations.
- We cannot strictly rank groups by mean happiness.

# *Univariate distributions: the measurement problem*

- More recently, Bond and Lang (2014) arrive at the same conclusion in a slightly different way.
- If the latent distribution can be represented by a a two-parameter distribution (e.g. cdf being a function of mean and variance) with different parameters for each group being compared - like in any estimation using ordered probit or ordered logit - then the CDFs of two groups will always cross except for a zero-probability event where the variances are the same, so there cannot be FOD which is necessary for the mean of two distributions to be invariant to monotonic transformations.
- We cannot strictly rank groups by mean happiness.

- More recently, Bond and Lang (2014) arrive at the same conclusion in a slightly different way.
- If the latent distribution can be represented by a a two-parameter distribution (e.g. cdf being a function of mean and variance) with different parameters for each group being compared - like in any estimation using ordered probit or ordered logit - then the CDFs of two groups will always cross except for a zero-probability event where the variances are the same, so there cannot be FOD which is necessary for the mean of two distributions to be invariant to monotonic transformations.
- We cannot strictly rank groups by mean happiness.

# Univariate distributions: the measurement problem

- Key concepts of standard inequality measurement literature are difficult to interpret in an ordinal framework. It is difficult to imagine a meaningful version of a Pigou-Dalton Transfer in the health context. “Transferring” health (?) from a healthier to a less healthy individual may exacerbate income inequality if a healthier individual is also a poorer one.
- Solution: consider an underlying social good that is transferrable and affects the distribution of health status e.g. access to health care.
- But then we need to know whether a transfer is sufficient to change the distribution of health status. So again, without further assumptions on the underlying variable, the impact of a Pigou-Dalton transfer will remain indeterminate.
- One solution: to develop a theory that deals directly with distributions of ordinal variables.

- Key concepts of standard inequality measurement literature are difficult to interpret in an ordinal framework. It is difficult to imagine a meaningful version of a Pigou-Dalton Transfer in the health context. “Transferring” health (?) from a healthier to a less healthy individual may exacerbate income inequality if a healthier individual is also a poorer one.
- Solution: consider an underlying social good that is transferrable and affects the distribution of health status e.g. access to health care.
- But then we need to know whether a transfer is sufficient to change the distribution of health status. So again, without further assumptions on the underlying variable, the impact of a Pigou-Dalton transfer will remain indeterminate.
- One solution: to develop a theory that deals directly with distributions of ordinal variables.

- Key concepts of standard inequality measurement literature are difficult to interpret in an ordinal framework. It is difficult to imagine a meaningful version of a Pigou-Dalton Transfer in the health context. “Transferring” health (?) from a healthier to a less healthy individual may exacerbate income inequality if a healthier individual is also a poorer one.
- Solution: consider an underlying social good that is transferrable and affects the distribution of health status e.g. access to health care.
- But then we need to know whether a transfer is sufficient to change the distribution of health status. So again, without further assumptions on the underlying variable, the impact of a Pigou-Dalton transfer will remain indeterminate.
- One solution: to develop a theory that deals directly with distributions of ordinal variables.

- Key concepts of standard inequality measurement literature are difficult to interpret in an ordinal framework. It is difficult to imagine a meaningful version of a Pigou-Dalton Transfer in the health context. "Transferring" health (?) from a healthier to a less healthy individual may exacerbate income inequality if a healthier individual is also a poorer one.
- Solution: consider an underlying social good that is transferrable and affects the distribution of health status e.g. access to health care.
- But then we need to know whether a transfer is sufficient to change the distribution of health status. So again, without further assumptions on the underlying variable, the impact of a Pigou-Dalton transfer will remain indeterminate.
- One solution: to develop a theory that deals directly with distributions of ordinal variables.

- In standard inequality measurement literature, a *perfectly equal distribution* is the one in which every individual possesses the same amount of income which then by definition is mean income. Inequality indicates a deviation from such a perfectly equal distribution and an inequality index measures the degree of the deviation.
- Similarly, in an ordinal data framework inequality can be thought of as a distance from the perfectly equal distribution. A natural candidate for a perfectly equal distribution is a distribution for which every individual is in the same category. Yet while in a standard framework a perfectly equal distribution is unique with ordinal data there are as many perfectly equal distributions as there are categories.

- In standard inequality measurement literature, a *perfectly equal distribution* is the one in which every individual possesses the same amount of income which then by definition is mean income. Inequality indicates a deviation from such a perfectly equal distribution and an inequality index measures the degree of the deviation.
- Similarly, in an ordinal data framework inequality can be thought of as a distance from the perfectly equal distribution. A natural candidate for a perfectly equal distribution is a distribution for which every individual is in the same category. Yet while in a standard framework a perfectly equal distribution is unique with ordinal data there are as many perfectly equal distributions as there are categories.

- The way this problem has been dealt with so far in the literature is to resort to the notion of a *perfectly unequal distribution*. For many researchers the most natural candidate for a perfectly unequal distribution is the one where half of probability mass is concentrated in the lowest category and half of probability mass is concentrated in the highest category (Leik (1966), Berry and Mielke (1992a), Blair and Lacy (2000), Allison and Foster (2004), and Abul Naga and Yalcin (2008) and others). Such distribution has the advantage of being uniquely defined.
- Zheng (2008) argues that it is more polarization than inequality and it is hard to disagree ...

- The way this problem has been dealt with so far in the literature is to resort to the notion of a *perfectly unequal distribution*. For many researchers the most natural candidate for a perfectly unequal distribution is the one where half of probability mass is concentrated in the lowest category and half of probability mass is concentrated in the highest category (Leik (1966), Berry and Mielke (1992a), Blair and Lacy (2000), Allison and Foster (2004), and Abul Naga and Yalcin (2008) and others). Such distribution has the advantage of being uniquely defined.
- Zheng (2008) argues that it is more polarization than inequality and it is hard to disagree ...

- Apouey and Silber (2013) claim that maximal inequality obtains if all individuals except one are in the lowest category, whereas one individual is in the highest category.
- On the other hand, they claim that there is no inequality (and polarization) if all mass is concentrated in one cell. –> discontinuity problem
- Recently, Gravel, Magdalou and Moyes (2015) propose a relation which can be thought of as measuring inequality rather than bi-polarization. They relate so called Hammond transfers to a dominance curve $F_i := 2F_{i-1} + f_i$.

- Apouey and Silber (2013) claim that maximal inequality obtains if all individuals except one are in the lowest category, whereas one individual is in the highest category.
- On the other hand, they claim that there is no inequality (and polarization) if all mass is concentrated in one cell. –> discontinuity problem
- Recently, Gravel, Magdalou and Moyes (2015) propose a relation which can be thought of as measuring inequality rather than bi-polarization. They relate so called Hammond transfers to a dominance curve $F_i := 2F_{i-1} + f_i$.

- Apouey and Silber (2013) claim that maximal inequality obtains if all individuals except one are in the lowest category, whereas one individual is in the highest category.
- On the other hand, they claim that there is no inequality (and polarization) if all mass is concentrated in one cell. –> discontinuity problem
- Recently, Gravel, Magdalou and Moyes (2015) propose a relation which can be thought of as measuring inequality rather than bi-polarization. They relate so called Hammond transfers to a dominance curve $F_i := 2F_{i-1} + f_i$.

- We typically work with the following framework. A numerical representation of categories of ordinal variables $\mathbb{I} := \mathcal{I}_1 \times \mathcal{I}_2 = \{0, \dots, n\} \times \{0, \dots, m\}$ which is arbitrary as long as it preserves the ordering.
- Let $f$ be a probability distribution on the set $\mathbb{I}$ (independent of scale). Obviously we require $\sum_{i=0}^{n} \sum_{j=0}^{m} f_{ij} = 1$ and for all $(i, j) \in \mathbb{I}$, $f_{ij} \geq 0$. We define marginal distributions by $f_i^1 := \sum_{j=0}^{m} f_{ij}, \ f_j^2 := \sum_{i=0}^{n} f_{ij}$ and cumulative distributions by $F_i^1 := \sum_{k=0}^{i} f_k, \ F_j^2 := \sum_{l=0}^{j} f_l$. When speaking about univariate distribution let $f, g$ denote univariate distributions (i.e. $f^j, g^j$), and $F, G$ univariate cdfs, respectively.
- A multidimensional cumulative distribution function $F$ at $(i, j)$ equals $F_{ij} := \sum_{k=0}^{i} \sum_{l=0}^{j} f_{kl}$.
- For each dimension $j$ we define a median $m_j$ which is the number for which $F^j_{m_j - 1} \leq 1/2$ and $F^j_{m_j} \geq 1/2$. Finally, let inequality index be denoted by $I : \Lambda \mapsto \mathbb{R}$.

- We typically work with the following framework. A numerical representation of categories of ordinal variables $\mathbb{I} := \mathcal{I}_1 \times \mathcal{I}_2 = \{0, \ldots, n\} \times \{0, \ldots, m\}$ which is arbitrary as long as it preserves the ordering.
- Let $f$ be a probability distribution on the set $\mathbb{I}$ (independent of scale). Obviously we require $\sum_{i=0}^{n} \sum_{j=0}^{m} f_{ij} = 1$ and for all $(i,j) \in \mathbb{I}$, $f_{ij} \geq 0$. We define marginal distributions by $f_i^1 := \sum_{j=0}^{m} f_{ij}, \;\; f_j^2 := \sum_{i=0}^{n} f_{ij}$ and cumulative distributions by $F_i^1 := \sum_{k=0}^{i} f_k, \;\; F_j^2 := \sum_{l=0}^{j} f_l$. When speaking about univariate distribution let $f, g$ denote univariate distributions (i.e. $f^j, g^j$), and $F, G$ univariate cdfs, respectively.
- A multidimensional cumulative distribution function $F$ at $(i,j)$ equals $F_{ij} := \sum_{k=0}^{i} \sum_{l=0}^{j} f_{kl}$.
- For each dimension $j$ we define a median $m_j$ which is the number for which $F^j_{m_j - 1} \leq 1/2$ and $F^j_{m_j} \geq 1/2$. Finally, let inequality index be denoted by $I : \Lambda \mapsto \mathbb{R}$.

# *Univariate distributions: new theory for ordinal data*

- We typically work with the following framework. A numerical representation of categories of ordinal variables $\mathbb{I} := \mathcal{I}_1 \times \mathcal{I}_2 = \{0, \ldots, n\} \times \{0, \ldots, m\}$ which is arbitrary as long as it preserves the ordering.
- Let $f$ be a probability distribution on the set $\mathbb{I}$ (independent of scale). Obviously we require $\sum_{i=0}^{n} \sum_{j=0}^{m} f_{ij} = 1$ and for all $(i, j) \in \mathbb{I}$, $f_{ij} \geq 0$. We define marginal distributions by $f_i^1 := \sum_{j=0}^{m} f_{ij}, \;\; f_j^2 := \sum_{i=0}^{n} f_{ij}$ and cumulative distributions by $F_i^1 := \sum_{k=0}^{i} f_k, \;\; F_j^2 := \sum_{l=0}^{j} f_l$. When speaking about univariate distribution let $f, g$ denote univariate distributions (i.e. $f^j, g^j$), and $F, G$ univariate cdfs, respectively.
- A multidimensional cumulative distribution function $F$ at $(i, j)$ equals $F_{ij} := \sum_{k=0}^{i} \sum_{l=0}^{j} f_{kl}$.
- For each dimension $j$ we define a median $m_j$ which is the number for which $F^j_{m_j - 1} \leq 1/2$ and $F^j_{m_j} \geq 1/2$. Finally, let inequality index be denoted by $I : \Lambda \mapsto \mathbb{R}$.

# *Univariate distributions: new theory for ordinal data*

- We typically work with the following framework. A numerical representation of categories of ordinal variables $\mathbb{I} := \mathcal{I}_1 \times \mathcal{I}_2 = \{0, \ldots, n\} \times \{0, \ldots, m\}$ which is arbitrary as long as it preserves the ordering.
- Let $f$ be a probability distribution on the set $\mathbb{I}$ (independent of scale). Obviously we require $\sum_{i=0}^{n} \sum_{j=0}^{m} f_{ij} = 1$ and for all $(i,j) \in \mathbb{I}$, $f_{ij} \geq 0$. We define marginal distributions by $f_i^1 := \sum_{j=0}^{m} f_{ij}, \;\; f_j^2 := \sum_{i=0}^{n} f_{ij}$ and cumulative distributions by $F_i^1 := \sum_{k=0}^{i} f_k, \;\; F_j^2 := \sum_{l=0}^{j} f_l$. When speaking about univariate distribution let $f, g$ denote univariate distributions (i.e. $f^j, g^j$), and $F, G$ univariate cdfs, respectively.
- A multidimensional cumulative distribution function $F$ at $(i,j)$ equals $F_{ij} := \sum_{k=0}^{i} \sum_{l=0}^{j} f_{kl}$.
- For each dimension $j$ we define a median $m_j$ which is the number for which $F^j_{m_j - 1} \leq 1/2$ and $F^j_{m_j} \geq 1/2$. Finally, let inequality index be denoted by $I : \Lambda \mapsto \mathbb{R}$.

*Definition*

***Hammond transfer*** *We say that g is obtained from f via a Hammond transfer if there exist categories* $1 < l < s \leq m < j < k$ *such that* $g_i = f_i \, \forall_{i \neq l,s,m,j}$ *and*

$$g_l = f_l - \eta_1 \quad g_s = f_s + \eta_1$$

$$g_m = f_m + \eta_2 \quad g_j = f_j - \eta_2,$$

*where* $\eta_1, \eta_2 > 0$*.*

Gravel, Magdalou and Moyes (2015) define it by moving individuals between categories. The spread of the distribution of an ordinal attribute is reduced irrespective of whether the “gain” of one individual is equal to the “loss” of the other.

*Definition*

***FOD*** *Fixing $n \geq 1$ and allowing $f, g$ to be two probability distributions on $\mathcal{I}$.*

$$F \leqslant_{FOD} G \Leftrightarrow F_i \leq G_i \quad \text{for any } i.$$

*Theorem*

*(Allison and Foster 2004) Let $c = \{c_1^1 \leq \ldots \leq c_n^1\}$ be a scale on $\mathcal{I}$. Further, let $\mu_F := \sum_{i=1}^{n} f_i c_i, \mu_G$ denote the mean of $F, G$. $F$ first order dominates $G$ if and only if $\mu_F \geq \mu_G$ for every $c$.*

- Allison and Foster (2004) postulate that inequality increases when probability mass is moved away from the median. They introduce the following partial ordering $\leqslant_{AF}$. Mendelson (1987) proposes similar relation for any quintile.

**Definition**

***Unidimensional AF***

*Fixing $n \geq 1$ and allowing $f, g$ to be two probability distributions on $\mathcal{I}$.*

$$F \leqslant_{AF} F \Leftrightarrow$$

(AF1) *$F, G$ have a unique and common median $m$,*

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

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

- Interpretation of this ordering is intuitive, in particular, $F \leqslant_{AF} G$ when $F$ is more concentrated around the median than $G$.

- Allison and Foster (2004) postulate that inequality increases when probability mass is moved away from the median. They introduce the following partial ordering $\leqslant_{AF}$. Mendelson (1987) proposes similar relation for any quintile.

*Definition*

***Unidimensional AF***

*Fixing $n \geq 1$ and allowing $f, g$ to be two probability distributions on $\mathcal{I}$.*

$$F \leqslant_{AF} F \Leftrightarrow$$

(AF1) $F, G$ *have a unique and common median* $m$,

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

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

- Interpretation of this ordering is intuitive, in particular, $F \leqslant_{AF} G$ when $F$ is more concentrated around the median than $G$.

- Abul Naga and Yalcin (2008), and Lazar and Silber (2013) characterize indices based on AF relation.
- In the general case, $I$ is of the form: $I(f) = \frac{\psi(F)-\psi(\check{F})}{\psi(\hat{F})-\psi(\check{F})}$, where $\psi(F) = \sum_{i<m} h_1(F_i) - \sum_{i\geq m} h_2(F_i)$, and $h_1, h_2$ are increasing functions.
- They derive several classes of indices.

- Abul Naga and Yalcin (2008), and Lazar and Silber (2013) characterize indices based on AF relation.
- In the general case, $I$ is of the form: $I(f) = \frac{\psi(F)-\psi(\check{F})}{\psi(\hat{F})-\psi(\check{F})}$, where $\psi(F) = \sum_{i<m} h_1(F_i) - \sum_{i\geq m} h_2(F_i)$, and $h_1, h_2$ are increasing functions.
- They derive several classes of indices.

- Abul Naga and Yalcin (2008), and Lazar and Silber (2013) characterize indices based on AF relation.
- In the general case, $I$ is of the form: $I(f) = \frac{\psi(F)-\psi(\check{F})}{\psi(\hat{F})-\psi(\check{F})}$, where $\psi(F) = \sum_{i<m} h_1(F_i) - \sum_{i \geq m} h_2(F_i)$, and $h_1, h_2$ are increasing functions.
- They derive several classes of indices.

- Kobus (2015) proves an Atkinson - type theorem (Atkinson, 1970) for ordinal data.

**Definition**

***Median-preserving spread principle*** *An index does not decrease following a median-preserving spread i.e. a transfer of probability mass away from the median (in both directions).*

**Definition**

*A function is $T-convex$ if it does not decrease following a multiplication via a $T-convex$ matrix.*

- In short, $T-convex$ functions are matrix representations of median-preserving spreads.

- Kobus (2015) proves an Atkinson - type theorem (Atkinson, 1970) for ordinal data.

**Definition**

***Median-preserving spread principle*** *An index does not decrease following a median-preserving spread i.e. a transfer of probability mass away from the median (in both directions).*

**Definition**

*A function is $T-convex$ if it does not decrease following a multiplication via a $T-convex$ matrix.*

- In short, $T-convex$ functions are matrix representations of median-preserving spreads.

**Theorem**

*(Kobus 2015) The following conditions are equivalent.*

1. $f \leqslant_{AF} g$.
2. *g can be obtained from f via a finite sequence of median-preserving spreads.*
3. $I(f) \leq I(g)$ *for all indices satisfying 4.*
4. $I(f) \leq I(g)$ *for all* $T-convex$ *indices.*

- Kobus and Miłoś (2012) characterize indices decomposable by population subgroups (similarly to Shorrocks 1980, 1984 for cardinal data).
- DECOMP requires that an index be presented as some function of inequality values in subgroups and subgroup sizes expressed in percentages. For example, let $f = (0.25, 0.25, 0.50)$; $g = (0.30, 0.40, 0.30)$ and $\alpha = 0.5$. Total distribution is $0.5f + 0.5g = (0.275, 0.325, 0.40)$. Then, if the inequality index fulfills DECOMP the inequality value associated with the distribution $(0.275, 0.325, 0.40)$ can be decomposed into inequality values in groups $f$ and $g$.

- Kobus and Miłoś (2012) characterize indices decomposable by population subgroups (similarly to Shorrocks 1980, 1984 for cardinal data).
- DECOMP requires that an index be presented as some function of inequality values in subgroups and subgroup sizes expressed in percentages. For example, let $f = (0.25, 0.25, 0.50)$; $g = (0.30, 0.40, 0.30)$ and $\alpha = 0.5$. Total distribution is $0.5f + 0.5g = (0.275, 0.325, 0.40)$. Then, if the inequality index fulfills DECOMP the inequality value associated with the distribution $(0.275, 0.325, 0.40)$ can be decomposed into inequality values in groups $f$ and $g$.

**Theorem**

*(Kobus and Miłoś 2012) I is continuous and fulfills DECOMP if and only if I is of the form*

$$I(f) = h(\sum_{i=1}^{n} a_i f_i),$$

*where* $(a_1, a_2, \ldots, a_n) \in \mathbb{R}^n$, $h : \mathbb{R} \mapsto [0, 1]$ *is a continuous strictly monotonic function. Moreover, let I be decomposable. Then, I is consistent with AF relation if and only if either h is a strictly increasing function and* $a_i \geq a_{i+1}$ *when* $i < m$ *and* $a_i \leq a_{i+1}$ *when* $i \geq m$ *or h is a strictly decreasing function and* $a_i \leq a_{i+1}$ *when* $i < m$ *and* $a_i \geq a_{i+1}$ *when* $i \geq m$.

# *Univariate distributions: measures of inequality/polarization*

- Several measures have been proposed in the literature (Blair and Lacy 2000; Allison and Foster 2004; Apouey 2007; Abul Naga and Yalcin 2008). We focus on measures proposed in Abul Naga and Yalcin (2008).
- Let $n$ be the number of categories and $m$ denote the median.

$$I^{a,b}(f) = \frac{a\sum_{i<m} F_i - b\sum_{i\geq m} F_i + b(n+1-m)}{(a(m-1)+b(n-m))/2}; \quad a, b \geq 0.$$

- 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. Kobus and Miłoś (2012) prove that this is the only class decomposable by population subgroups (among measures proposed in the literature).

# *Univariate distributions: measures of inequality/polarization*

- Several measures have been proposed in the literature (Blair and Lacy 2000; Allison and Foster 2004; Apouey 2007; Abul Naga and Yalcin 2008). We focus on measures proposed in Abul Naga and Yalcin (2008).
- Let $n$ be the number of categories and $m$ denote the median.

$$I^{a,b}(f) = \frac{a\sum_{i<m} F_i - b\sum_{i\geq m} F_i + b(n+1-m)}{\left(a(m-1) + b(n-m)\right)/2}; \quad a, b \geq 0.$$

- 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. Kobus and Miłoś (2012) prove that this is the only class decomposable by population subgroups (among measures proposed in the literature).

# *Univariate distributions: measures of inequality/polarization*

- Several measures have been proposed in the literature (Blair and Lacy 2000; Allison and Foster 2004; Apouey 2007; Abul Naga and Yalcin 2008). We focus on measures proposed in Abul Naga and Yalcin (2008).
- Let $n$ be the number of categories and $m$ denote the median.

$$I^{a,b}(f) = \frac{a\sum_{i<m} F_i - b\sum_{i\geq m} F_i + b(n+1-m)}{\left(a(m-1)+b(n-m)\right)/2}; \;\; a,b \geq 0.$$

- 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. Kobus and Miłoś (2012) prove that this is the only class decomposable by population subgroups (among measures proposed in the literature).

# *Univariate distributions: measures of inequality/polarization*

- $$I_{\alpha,\beta}(f) = \frac{\sum_{i<m} F_i^{\alpha} - \sum_{i\geq m} F_i^{\beta} + (n+1-m)}{k_{\alpha,\beta} + (n+1-m)},$$

  where $k_{\alpha,\beta} = (m-1)\left(\frac{1}{2}\right)^{\alpha} - \left(1 + (n-m)\left(\frac{1}{2}\right)^{\beta}\right)$ and $\alpha, \beta \geq 1$.
- For a given value of $\beta$, the index becomes more sensitive to inequality in the lower end as $\alpha \to 1$ and in the higher end as $\alpha \to \infty$. Conversely, for given value of $\alpha$, the index becomes more sensitive to inequality in the top of the distribution as $\beta$ rises.
- Recently, Abul Naga and Stapengurst (2015) provide estimation for both $I_{\alpha,\beta}$ and $I^{a,b}$.

# *Univariate distributions: measures of inequality/polarization*

- $$I_{\alpha,\beta}(f) = \frac{\sum_{i<m} F_i^{\alpha} - \sum_{i\geq m} F_i^{\beta} + (n+1-m)}{k_{\alpha,\beta} + (n+1-m)},$$

  where $k_{\alpha,\beta} = (m-1)\left(\frac{1}{2}\right)^{\alpha} - \left(1 + (n-m)\left(\frac{1}{2}\right)^{\beta}\right)$ and $\alpha, \beta \geq 1$.
- For a given value of $\beta$, the index becomes more sensitive to inequality in the lower end as $\alpha \to 1$ and in the higher end as $\alpha \to \infty$. Conversely, for given value of $\alpha$, the index becomes more sensitive to inequality in the top of the distribution as $\beta$ rises.
- Recently, Abul Naga and Stapengurst (2015) provide estimation for both $I_{\alpha,\beta}$ and $I^{a,b}$.

# *Univariate distributions: measures of inequality/polarization*

- $$I_{\alpha,\beta}(f) = \frac{\sum_{i<m} F_i^{\alpha} - \sum_{i\geq m} F_i^{\beta} + (n+1-m)}{k_{\alpha,\beta} + (n+1-m)},$$
  where $k_{\alpha,\beta} = (m-1)\left(\frac{1}{2}\right)^{\alpha} - \left(1 + (n-m)\left(\frac{1}{2}\right)^{\beta}\right)$ and $\alpha, \beta \geq 1$.
- For a given value of $\beta$, the index becomes more sensitive to inequality in the lower end as $\alpha \to 1$ and in the higher end as $\alpha \to \infty$. Conversely, for given value of $\alpha$, the index becomes more sensitive to inequality in the top of the distribution as $\beta$ rises.
- Recently, Abul Naga and Stapengurst (2015) provide estimation for both $I_{\alpha,\beta}$ and $I^{a,b}$.

- Cowell and Flachaire (2015) develop a class of inequality indices conditional on a reference point. In particular, they characterize the following family of indices

$$I_\alpha(s, e) := \frac{1}{\alpha(\alpha - 1)} \left( \frac{1}{n} \sum_{i=1}^{n} s_i^\alpha - e^\alpha \right),$$

where $s_i$ is person $i$ status (which can be measured by person $i$'s position in the distribution of an ordinal variable e.g. downward looking status - the proportion of those below me), $e$ is the reference point and $\alpha$ measures sensitivity to inequality in different parts of distributions. Also $e$ can be a function of $s$.

- Cowell and Flachaire (2015) consider different reference points and choose maximal status ($e = 1$). This is the only situation where for downward or upward looking status distance from the reference point is zero for everyone.
- In the case where $e = 1$ they obtain

$$I_\alpha(s, 1) := \frac{1}{\alpha(\alpha - 1)} \left( \frac{1}{n} \sum_{i=1}^{n} s_i^\alpha - 1 \right), \quad \alpha \neq 0, 1$$

$$I_\alpha(s, 1) := -\frac{1}{n} \sum_{i=1}^{n} \ln s_i, \quad \alpha = 0.$$

- They develop statistical properties of $I_\alpha$.

- Cowell and Flachaire (2015) consider different reference points and choose maximal status ($e = 1$). This is the only situation where for downward or upward looking status distance from the reference point is zero for everyone.
- In the case where $e = 1$ they obtain

$$I_\alpha(s, 1) := \frac{1}{\alpha(\alpha - 1)} \left( \frac{1}{n} \sum_{i=1}^{n} s_i^\alpha - 1 \right), \quad \alpha \neq 0, 1$$

$$I_\alpha(s, 1) := -\frac{1}{n} \sum_{i=1}^{n} \ln s_i, \quad \alpha = 0.$$

- They develop statistical properties of $I_\alpha$.

- Cowell and Flachaire (2015) consider different reference points and choose maximal status ($e = 1$). This is the only situation where for downward or upward looking status distance from the reference point is zero for everyone.
- In the case where $e = 1$ they obtain

$$I_\alpha(s,1) := \frac{1}{\alpha(\alpha-1)}\left(\frac{1}{n}\sum_{i=1}^{n} s_i^\alpha - 1\right), \quad \alpha \neq 0,1$$

$$I_\alpha(s,1) := -\frac{1}{n}\sum_{i=1}^{n} \ln s_i, \quad \alpha = 0.$$

- They develop statistical properties of $I_\alpha$.

- Lv et al. (2015) characterize the following family of inequality measures.

$$I(f) := \sum_{i=1}^{n} \sum_{j \neq i} h(|i-j|) f_i f_j,$$

where $h : \{0, \ldots, n-1\} \mapsto (0, \infty)$ is an increasing function.

- The degree of inequality between any two categories $i, j$ is captured, and $h$ increases with the distance between categories.
- Depending on $h$, this class includes various measures.

- Lv et al. (2015) characterize the following family of inequality measures.

$$I(f) := \sum_{i=1}^{n} \sum_{j \neq i} h(|i - j|) f_i f_j,$$

where $h : \{0, \ldots, n-1\} \mapsto (0, \infty)$ is an increasing function.

- The degree of inequality between any two categories $i, j$ is captured, and $h$ increases with the distance between categories.
- Depending on $h$, this class includes various measures.

- Lv et al. (2015) characterize the following family of inequality measures.

$$I(f) := \sum_{i=1}^{n} \sum_{j \neq i} h(|i-j|) f_i f_j,$$

where $h : \{0, \ldots, n-1\} \mapsto (0, \infty)$ is an increasing function.

- The degree of inequality between any two categories $i, j$ is captured, and $h$ increases with the distance between categories.
- Depending on $h$, this class includes various measures.

- As mentioned, bivariate distributions of ordinal data emerge most often in the so called socioeconomic inequalities in health literature.
- Typically to assess income-related inequality in health we use concentration curve and a related concentration index (Wagstaff et al. 1991; Kakwani et al. 1997) which plots the cumulative percentage of health variable vs cumulative percentage of the sample, ranked by incomes, from the poorest to the richest.
- For ordinal data common problems with scaling apply i.e. the use of ordinal variables to compute an index developed for ratio-scale variables (Makdisi and Yazbeck 2014).
- Often regression coefficient is interpreted as the gradient. This picks up linear dependence.

# Bivariate distributions

- As mentioned, bivariate distributions of ordinal data emerge most often in the so called socioeconomic inequalities in health literature.
- Typically to assess income-related inequality in health we use concentration curve and a related concentration index (Wagstaff et al. 1991; Kakwani et al. 1997) which plots the cumulative percentage of health variable vs cumulative percentage of the sample, ranked by incomes, from the poorest to the richest.
- For ordinal data common problems with scaling apply i.e. the use of ordinal variables to compute an index developed for ratio-scale variables (Makdisi and Yazbeck 2014).
- Often regression coefficient is interpreted as the gradient. This picks up linear dependence.

# *Bivariate distributions*

- As mentioned, bivariate distributions of ordinal data emerge most often in the so called socioeconomic inequalities in health literature.
- Typically to assess income-related inequality in health we use concentration curve and a related concentration index (Wagstaff et al. 1991; Kakwani et al. 1997) which plots the cumulative percentage of health variable vs cumulative percentage of the sample, ranked by incomes, from the poorest to the richest.
- For ordinal data common problems with scaling apply i.e. the use of ordinal variables to compute an index developed for ratio-scale variables (Makdisi and Yazbeck 2014).
- Often regression coefficient is interpreted as the gradient. This picks up linear dependence.

# *Bivariate distributions*

- As mentioned, bivariate distributions of ordinal data emerge most often in the so called socioeconomic inequalities in health literature.
- Typically to assess income-related inequality in health we use concentration curve and a related concentration index (Wagstaff et al. 1991; Kakwani et al. 1997) which plots the cumulative percentage of health variable vs cumulative percentage of the sample, ranked by incomes, from the poorest to the richest.
- For ordinal data common problems with scaling apply i.e. the use of ordinal variables to compute an index developed for ratio-scale variables (Makdisi and Yazbeck 2014).
- Often regression coefficient is interpreted as the gradient. This picks up linear dependence.

Zheng (2011) proposes the following definition of Generalized Lorenz dominance for welfare comparisons.

$$GL(f, l) := \sum_{i=1}^{l} \sum_{j=1}^{m} \pi_i f_{ij} h_j, \quad l = 1, 2, \ldots, n,$$

where $\pi = (\pi_i)_{i=1}^{n}$ is a socioeconomic structure i.e. income quintiles, and $h_j$ is the value of health (scale –> ordered).

**Theorem**

*(Zheng 2011) For a given $\pi$ and for all increasing scales h, f Generalized Lorenz dominates g i.e.*

$$GL(f, l) \geq GL(g, l) \ \forall_l \iff \sum_{i=1}^{l} \sum_{j=1}^{k} \pi_i f_{ij} \leq \sum_{i=1}^{l} \sum_{j=1}^{k} \pi_i g_{ij} \ \forall_{k,l},$$

*where $l = 1, \ldots, n$ and $k = 1, \ldots, m$.*

# Bivariate distributions: dominance conditions

- For inequality comparisons, he defines relative and absolute Lorenz curves which are mean-normalizations of GL. The relative (absolute) Lorenz curve remans unchanged if all expected health levels are increased proportionally (by the same amount).
- GL dominance is related to welfare functions which are monotone and satisfy Pigou-Dalton Transfer (PDT) axiom. Here PDT involves a transfer of health from a healthier class to a less healthy which is problematic.
- Zheng (2011) proposes to define welfare functions on the vector of expected health $e(f) := \left(\sum_{j=1}^{m} f_{1j}h_j, \ldots, \sum_{j=1}^{m} f_{nj}h_j\right)$ and to assume that this vector is in non-decreasing order i.e. higher socioeconomic class has better expected health.
- Then, he argues, PDT becomes more acceptable because a healthier class also has higher income.

# *Bivariate distributions: dominance conditions*

- For inequality comparisons, he defines relative and absolute Lorenz curves which are mean-normalizations of GL. The relative (absolute) Lorenz curve remans unchanged if all expected health levels are increased proportionally (by the same amount).
- GL dominance is related to welfare functions which are monotone and satisfy Pigou-Dalton Transfer (PDT) axiom. Here PDT involves a transfer of health from a healthier class to a less healthy which is problematic.
- Zheng (2011) proposes to define welfare functions on the vector of expected health $e(f) := \left(\sum_{j=1}^{m} f_{1j}h_j, \ldots, \sum_{j=1}^{m} f_{nj}h_j\right)$ and to assume that this vector is in non-decreasing order i.e. higher socioeconomic class has better expected health.
- Then, he argues, PDT becomes more acceptable because a healthier class also has higher income.

- For inequality comparisons, he defines relative and absolute Lorenz curves which are mean-normalizations of GL. The relative (absolute) Lorenz curve remans unchanged if all expected health levels are increased proportionally (by the same amount).
- GL dominance is related to welfare functions which are monotone and satisfy Pigou-Dalton Transfer (PDT) axiom. Here PDT involves a transfer of health from a healthier class to a less healthy which is problematic.
- Zheng (2011) proposes to define welfare functions on the vector of expected health $e(f) := \left(\sum_{j=1}^{m} f_{1j}h_j, \ldots, \sum_{j=1}^{m} f_{nj}h_j\right)$ and to assume that this vector is in non-decreasing order i.e. higher socioeconomic class has better expected health.
- Then, he argues, PDT becomes more acceptable because a healthier class also has higher income.

# Bivariate distributions: dominance conditions

- For inequality comparisons, he defines relative and absolute Lorenz curves which are mean-normalizations of GL. The relative (absolute) Lorenz curve remans unchanged if all expected health levels are increased proportionally (by the same amount).
- GL dominance is related to welfare functions which are monotone and satisfy Pigou-Dalton Transfer (PDT) axiom. Here PDT involves a transfer of health from a healthier class to a less healthy which is problematic.
- Zheng (2011) proposes to define welfare functions on the vector of expected health $e(f) := \left(\sum_{j=1}^{m} f_{1j}h_j, \ldots, \sum_{j=1}^{m} f_{nj}h_j\right)$ and to assume that this vector is in non-decreasing order i.e. higher socioeconomic class has better expected health.
- Then, he argues, PDT becomes more acceptable because a healthier class also has higher income.

# Bivariate distributions: dominance conditions

- The problem with Zheng (2011) is that $\pi$ appears ad hoc, without relation to $f$ and is assumed to be the same for distributions under comparison. The only empirical application is thus income-health matrix, where $\pi$ is income quintile distribution.
- One needs to modify Zheng's conditions to use it for the case of two ordinal variables e.g. education-health distribution (work in progress).
- Sonne-Schmidt et al. (2013) characterize all comparisons in the $2 \times 2$ case i.e. two binary variables. But there does not seem to be an easy method to generalize this case even a bit.
- There is a related literature in segregation (Reardon 2009, Cuhadaroglu 2015) and dissimilarity (Zoli and Andreoli 2015).

- The problem with Zheng (2011) is that $\pi$ appears ad hoc, without relation to $f$ and is assumed to be the same for distributions under comparison. The only empirical application is thus income-health matrix, where $\pi$ is income quintile distribution.
- One needs to modify Zheng's conditions to use it for the case of two ordinal variables e.g. education-health distribution (work in progress).
- Sonne-Schmidt et al. (2013) characterize all comparisons in the $2 \times 2$ case i.e. two binary variables. But there does not seem to be an easy method to generalize this case even a bit.
- There is a related literature in segregation (Reardon 2009, Cuhadaroglu 2015) and dissimilarity (Zoli and Andreoli 2015).

- The problem with Zheng (2011) is that $\pi$ appears ad hoc, without relation to $f$ and is assumed to be the same for distributions under comparison. The only empirical application is thus income-health matrix, where $\pi$ is income quintile distribution.
- One needs to modify Zheng's conditions to use it for the case of two ordinal variables e.g. education-health distribution (work in progress).
- Sonne-Schmidt et al. (2013) characterize all comparisons in the 2 $\times$ 2 case i.e. two binary variables. But there does not seem to be an easy method to generalize this case even a bit.
- There is a related literature in segregation (Reardon 2009, Cuhadaroglu 2015) and dissimilarity (Zoli and Andreoli 2015).

# *Bivariate distributions: dominance conditions*

- The problem with Zheng (2011) is that $\pi$ appears ad hoc, without relation to $f$ and is assumed to be the same for distributions under comparison. The only empirical application is thus income-health matrix, where $\pi$ is income quintile distribution.
- One needs to modify Zheng's conditions to use it for the case of two ordinal variables e.g. education-health distribution (work in progress).
- Sonne-Schmidt et al. (2013) characterize all comparisons in the $2 \times 2$ case i.e. two binary variables. But there does not seem to be an easy method to generalize this case even a bit.
- There is a related literature in segregation (Reardon 2009, Cuhadaroglu 2015) and dissimilarity (Zoli and Andreoli 2015).

# *Bivariate distributions: other solutions*

- Makdisi and Yazbeck (2014) omit the problem of cardinalizing health variable by changing the way in which one measures health.
- In particular, the number of health problems is a variable with a fixed measurement scale. Thus, they use the so called counting approach (Alkire and Foster 2010 in poverty measurement).
- They have to transform the available distribution into $0-1$ i.e. something counts as a problem if some specified threshold/cutt-off point is met. This causes substantial loss of information.

# *Bivariate distributions: other solutions*

- Makdisi and Yazbeck (2014) omit the problem of cardinalizing health variable by changing the way in which one measures health.
- In particular, the number of health problems is a variable with a fixed measurement scale. Thus, they use the so called counting approach (Alkire and Foster 2010 in poverty measurement).
- They have to transform the available distribution into $0-1$ i.e. something counts as a problem if some specified threshold/cutt-off point is met. This causes substantial loss of information.

# *Bivariate distributions: other solutions*

- Makdisi and Yazbeck (2014) omit the problem of cardinalizing health variable by changing the way in which one measures health.
- In particular, the number of health problems is a variable with a fixed measurement scale. Thus, they use the so called counting approach (Alkire and Foster 2010 in poverty measurement).
- They have to transform the available distribution into $0 - 1$ i.e. something counts as a problem if some specified threshold/cutt-off point is met. This causes substantial loss of information.

# *Bivariate distributions: measuring interdependence in ordinal data*

- Few authors (Decancq 2012, 2013; Meyer and Strulovici 2012, 2015; Kobus 2015) focus on measuring interdependence in ordinal data, which is a distinctive feature of multivariate distributions as opposed to univariate is interdependence between variables (i.e. ceteris paribus, higher association means higher inequality).
- Acknowledging association explicitly makes measures attribute decomposable (Naga and Geoffard 2006) by definition, which is a useful property.
- Given the problems with cardinalization, it seems more appropriate to measure association not between health and income/education/SES, but between the *distributions* of health and income/education/SES.

## *Bivariate distributions: measuring interdependence in ordinal data*

- Few authors (Decancq 2012, 2013; Meyer and Strulovici 2012, 2015; Kobus 2015) focus on measuring interdependence in ordinal data, which is a distinctive feature of multivariate distributions as opposed to univariate is interdependence between variables (i.e. ceteris paribus, higher association means higher inequality).
- Acknowledging association explicitly makes measures attribute decomposable (Naga and Geoffard 2006) by definition, which is a useful property.
- Given the problems with cardinalization, it seems more appropriate to measure association not between health and income/education/SES, but between the *distributions* of health and income/education/SES.

# *Bivariate distributions: measuring interdependence in ordinal data*

- Few authors (Decancq 2012, 2013; Meyer and Strulovici 2012, 2015; Kobus 2015) focus on measuring interdependence in ordinal data, which is a distinctive feature of multivariate distributions as opposed to univariate is interdependence between variables (i.e. ceteris paribus, higher association means higher inequality).
- Acknowledging association explicitly makes measures attribute decomposable (Naga and Geoffard 2006) by definition, which is a useful property.
- Given the problems with cardinalization, it seems more appropriate to measure association not between health and income/education/SES, but between the *distributions* of health and income/education/SES.

# *Bivariate distributions: measuring interdependence in ordinal data*

- There is a measure of association that does exactly this. It is a copula.
- Copulas are well-known in mathematics and statistics due to the celebrated Sklar's theorem (Sklar, 1959) which states that copula and marginal distributions characterize joint distribution fully. Copula is a bivariate function that binds marginal distributions and returns joint distribution.
- Formally, a bi-dimensional copula $C : [0,1]^2 \mapsto [0,1]$ is a function such that

$$F_{ij} = C\left(F_i^1, F_j^2\right).$$

# *Bivariate distributions: measuring interdependence in ordinal data*

- There is a measure of association that does exactly this. It is a copula.
- Copulas are well-known in mathematics and statistics due to the celebrated Sklar's theorem (Sklar, 1959) which states that copula and marginal distributions characterize joint distribution fully. Copula is a bivariate function that binds marginal distributions and returns joint distribution.
- Formally, a bi-dimensional copula $C : [0,1]^2 \mapsto [0,1]$ is a function such that

$$F_{ij} = C\left(F_i^1, F_j^2\right).$$

## *Bivariate distributions: measuring interdependence in ordinal data*

- There is a measure of association that does exactly this. It is a copula.
- Copulas are well-known in mathematics and statistics due to the celebrated Sklar's theorem (Sklar, 1959) which states that copula and marginal distributions characterize joint distribution fully. Copula is a bivariate function that binds marginal distributions and returns joint distribution.
- Formally, a bi-dimensional copula $C : [0,1]^2 \mapsto [0,1]$ is a function such that

$$F_{ij} = C\left(F_i^1, F_j^2\right).$$

- Copula is invariant under increasing transformations of variables, whereas cdf is not.
- Monotone transformation change the dependence structure. Copula is “designed” exactly for situations when such re-scaling should not change underlying distributional. As Schweizer and Wolff (1981) note, “it is precisely the copula which captures those properties of the joint distribution which are invariant under (...) strictly increasing transformations.”
- This is a very natural concept when studying mobility IF mobility is based on relative position in the income distribution - the probability that the child from bottom quintile gets to the top quintile.

- Copula is invariant under increasing transformations of variables, whereas cdf is not.
- Monotone transformation change the dependence structure. Copula is “designed” exactly for situations when such re-scaling should not change underlying distributional. As Schweizer and Wolff (1981) note, “it is precisely the copula which captures those properties of the joint distribution which are invariant under (...) strictly increasing transformations.”
- This is a very natural concept when studying mobility IF mobility is based on relative position in the income distribution - the probability that the child from bottom quintile gets to the top quintile.

- Copula is invariant under increasing transformations of variables, whereas cdf is not.
- Monotone transformation change the dependence structure. Copula is "designed" exactly for situations when such re-scaling should not change underlying distributional. As Schweizer and Wolff (1981) note, "it is precisely the copula which captures those properties of the joint distribution which are invariant under (...) strictly increasing transformations."
- This is a very natural concept when studying mobility IF mobility is based on relative position in the income distribution - the probability that the child from bottom quintile gets to the top quintile.

- Decancq (2013) proposes several measures of dependence based on a copula. His results are, however, for continuous distributions, and their application to discrete data is far from straightforward.
- In general, using copulas for discrete distributions is not easy. Genest and Neslehova (2007) describe all the bad things that happen with copulas for discrete data.
- Main problem: unidentifiability.

# *Bivariate distributions: measuring interdependence*

- Decancq (2013) proposes several measures of dependence based on a copula. His results are, however, for continuous distributions, and their application to discrete data is far from straightforward.
- In general, using copulas for discrete distributions is not easy. Genest and Neslehova (2007) describe all the bad things that happen with copulas for discrete data.
- Main problem: unidentifiability.

## *Bivariate distributions: measuring interdependence*

- Decancq (2013) proposes several measures of dependence based on a copula. His results are, however, for continuous distributions, and their application to discrete data is far from straightforward.
- In general, using copulas for discrete distributions is not easy. Genest and Neslehova (2007) describe all the bad things that happen with copulas for discrete data.
- Main problem: unidentifiability.

- Meyer and Strulovici (2015) are concerned with comparing distributions in terms of interdependence only (so margins are identical, and the unidentifiability does not cause too many problems). They define supermodular ordering which in this case is equivalent to an ordering on copulas used by Kobus (2015, work in progress).
- Yet in empirical comparisons we often have discrete distributions with different margins. Kobus (2015) is concerned with comparing such distributions. She deals with unidentifiability by considering the whole set of copulas associated with a given distribution and studying its bounds which were specified in mathematics literature by Carley (2002) and further by Neslehova (2007).
- This allows to define concordance (increasing dependence) ordering, which can be then combines with different unidimensional dominance conditions (e.g. AF dominance).

- Meyer and Strulovici (2015) are concerned with comparing distributions in terms of interdependence only (so margins are identical, and the unidentifiability does not cause too many problems). They define supermodular ordering which in this case is equivalent to an ordering on copulas used by Kobus (2015, work in progress).
- Yet in empirical comparisons we often have discrete distributions with different margins. Kobus (2015) is concerned with comparing such distributions. She deals with unidentifiability by considering the whole set of copulas associated with a given distribution and studying its bounds which were specified in mathematics literature by Carley (2002) and further by Neslehova (2007).
- This allows to define concordance (increasing dependence) ordering, which can be then combines with different unidimensional dominance conditions (e.g. AF dominance).

- Meyer and Strulovici (2015) are concerned with comparing distributions in terms of interdependence only (so margins are identical, and the unidentifiability does not cause too many problems). They define supermodular ordering which in this case is equivalent to an ordering on copulas used by Kobus (2015, work in progress).
- Yet in empirical comparisons we often have discrete distributions with different margins. Kobus (2015) is concerned with comparing such distributions. She deals with unidentifiability by considering the whole set of copulas associated with a given distribution and studying its bounds which were specified in mathematics literature by Carley (2002) and further by Neslehova (2007).
- This allows to define concordance (increasing dependence) ordering, which can be then combines with different unidimensional dominance conditions (e.g. AF dominance).

- Decancq (2012) provides a generalization of bivariate association concepts to trivariate distributions (elementary rearrangements).
- In three dimensions positive association no longer equals minus negative association. This very quickly gets very technical, but especially in health economics many health indicators are often available.
- Meyer and Strulovici (2012) provide a range of dependence orderings in several dimensions (and rank them!).

- Decancq (2012) provides a generalization of bivariate association concepts to trivariate distributions (elementary rearrangements).
- In three dimensions positive association no longer equals minus negative association. This very quickly gets very technical, but especially in health economics many health indicators are often available.
- Meyer and Strulovici (2012) provide a range of dependence orderings in several dimensions (and rank them!).

- Decancq (2012) provides a generalization of bivariate association concepts to trivariate distributions (elementary rearrangements).
- In three dimensions positive association no longer equals minus negative association. This very quickly gets very technical, but especially in health economics many health indicators are often available.
- Meyer and Strulovici (2012) provide a range of dependence orderings in several dimensions (and rank them!).

# *Summary: what has been done and can be used when measuring inequality in ordinal data*

- If one wants to measure inequality/polarization in health, happiness, educational attainment or other types of inequality when ordinal data are present, there is now a whole array of measures that can be used for which estimation has also been developed - Abul Naga and Yalcin (2008) two classes, with the extension proposed by Kobus and Miłoś (2012), or Cowell and Flachaire (2015).
- If one is interested in more robust, non-parametric comparisons, there are dominance conditions that can be used - Allison and Foster (2004) or Gravel, Magdalou and Moyes (2015).
- For income-related or socioeconomic-related inequalities in health, happiness etc. one may start with checking Zheng (2011) conditions.

# *Summary: what has been done and can be used when measuring inequality in ordinal data*

- If one wants to measure inequality/polarization in health, happiness, educational attainment or other types of inequality when ordinal data are present, there is now a whole array of measures that can be used for which estimation has also been developed - Abul Naga and Yalcin (2008) two classes, with the extension proposed by Kobus and Miłoś (2012), or Cowell and Flachaire (2015).
- If one is interested in more robust, non-parametric comparisons, there are dominance conditions that can be used - Allison and Foster (2004) or Gravel, Magdalou and Moyes (2015).
- For income-related or socioeconomic-related inequalities in health, happiness etc. one may start with checking Zheng (2011) conditions.

# *Summary: what has been done and can be used when measuring inequality in ordinal data*

- If one wants to measure inequality/polarization in health, happiness, educational attainment or other types of inequality when ordinal data are present, there is now a whole array of measures that can be used for which estimation has also been developed - Abul Naga and Yalcin (2008) two classes, with the extension proposed by Kobus and Miłoś (2012), or Cowell and Flachaire (2015).
- If one is interested in more robust, non-parametric comparisons, there are dominance conditions that can be used - Allison and Foster (2004) or Gravel, Magdalou and Moyes (2015).
- For income-related or socioeconomic-related inequalities in health, happiness etc. one may start with checking Zheng (2011) conditions.

# Summary: future research

- There is lack of normative inequality theory in the ordinal data context. In the standard literature, welfare functions can be expressed as functions of means ("efficiency") and inequality ("equity"). Here we do not have an appropriate "efficiency statistic".
- The literature on bivariate comparisons is small. There is a need for new dominance conditions, measures.
- There seems to be a need for developing modelling approaches which model the association structure more flexibly so that the ordinal nature of the data is accounted for.
- The comparisons of multivariate distributions remain to be solved.

# *Summary: future research*

- There is lack of normative inequality theory in the ordinal data context. In the standard literature, welfare functions can be expressed as functions of means ("efficiency") and inequality ("equity"). Here we do not have an appropriate "efficiency statistic".
- The literature on bivariate comparisons is small. There is a need for new dominance conditions, measures.
- There seems to be a need for developing modelling approaches which model the association structure more flexibly so that the ordinal nature of the data is accounted for.
- The comparisons of multivariate distributions remain to be solved.

- There is lack of normative inequality theory in the ordinal data context. In the standard literature, welfare functions can be expressed as functions of means ("efficiency") and inequality ("equity"). Here we do not have an appropriate "efficiency statistic".
- The literature on bivariate comparisons is small. There is a need for new dominance conditions, measures.
- There seems to be a need for developing modelling approaches which model the association structure more flexibly so that the ordinal nature of the data is accounted for.
- The comparisons of multivariate distributions remain to be solved.

# Summary: future research

- There is lack of normative inequality theory in the ordinal data context. In the standard literature, welfare functions can be expressed as functions of means (“efficiency”) and inequality (“equity”). Here we do not have an appropriate “efficiency statistic”.
- The literature on bivariate comparisons is small. There is a need for new dominance conditions, measures.
- There seems to be a need for developing modelling approaches which model the association structure more flexibly so that the ordinal nature of the data is accounted for.
- The comparisons of multivariate distributions remain to be solved.