DISCUSSION PAPERS IN STATISTICS AND ECONOMETRICS

SEMINAR OF ECONOMIC AND SOCIAL STATISTICS
UNIVERSITY OF COLOGNE

No. 3/08

Measuring Polarization via Poverty and Affluence

by

Christoph Scheicher

2nd version
June 24, 2010

DISKUSSIONSBEITRÄGE ZUR STATISTIK UND ÖKONOMETRIE

SEMINAR FÜR WIRTSCHAFTS- UND SOZIALSTATISTIK
UNIVERSITÄT ZU KÖLN

Albertus-Magnus-Platz, D-50923 Köln, Deutschland

This page intentionally left blank.

DISCUSSION PAPERS IN STATISTICS AND ECONOMETRICS

SEMINAR OF ECONOMIC AND SOCIAL STATISTICS
UNIVERSITY OF COLOGNE

No. 3/08

Measuring Polarization via Poverty and Affluence

by

Christoph Scheicher[1]

2nd version
June 24, 2010

Abstract

The decline of the middle class has been investigated as a principal aspect of social polarization (Wolfson 1994, 1997). Wang and Tsui 2000 have characterized a class of polarization measures by postulates on normalization, increasing spread and increasing bipolarity. The present paper generalizes this class of measures. It defines polarization by aggregating measures of poverty and of affluence, focussing on incomes outside a middle class interval. The approach is applied to German data on income distribution.

JEL: D31, D63, I32.

Keywords: decline of middle class, income distribution, income richness.

[1]Department of Economic and Social Statistics, Statistics & Econometrics, University of Cologne, Albertus-Magnus-Platz, 50923 Cologne, Germany.

This page intentionally left blank.

Measuring Polarization via Poverty and Affluence

Christoph Scheicher*

June 24, 2010

Abstract[1]

The decline of the middle class has been investigated as a principal aspect of social polarization (Wolfson 1994, 1997). Wang and Tsui (2000) have characterized a class of polarization measures by postulates on normalization, increasing spread and increasing bipolarity. The present paper generalizes this class of measures. It defines polarization by aggregating measures of poverty and of affluence, focussing on incomes outside a middle class interval. The approach is applied to German data on income distribution.

JEL: D31, D63, I32

Keywords: decline of middle class, income distribution, income richness

1 Introduction

The concept of social polarization is used to describe how seriously different groups of a society are divided.

One approach to measure polarization was developed by Wolfson (1994, 1997), focussing on the "decline of middle class". Wang and Tsui (2000) follow the Wolfson approach by defining indices of polarization. They consider distances from a

*Universität zu Köln, Seminar für Wirtschafts- und Sozialstatistik, Albertus-Magnus-Platz, 50923 Köln, scheicher@statistik.uni-koeln.de

[1]I thank Karl Mosler and Rainer Dyckerhoff for reading the paper and providing many valuable comments.

central point of the income distribution, the median income. They transform these distances by an increasing function and measure polarization by a mean of transformed distances.

The other – even more popular – approach is introduced by Esteban and Ray (1994): Population is divided somehow in several groups. A population is heavily polarized, if firstly the population is divided in few groups, secondly the interesting variable, e.g. income has very small spread within each group, but there are enormous differences between those groups. Thirdly all groups have to be sufficiently large. Many authors followed and modified the Esteban and Ray (1994) approach, e.g. Chakravarty and Majumder (2001), D´Ambrosio (2001), Duclos et al. (2004) and Esteban et al. (2007).

Nonetheless we focus on the Wolfson approach and modify Wang and Tsui (2000) polarization measure by using a middle class interval instead of a central middle class point and calculate the distances to this middle class interval.[2] We use different functions to transform the distances, since this allows us to treat the poor differently from the rich. We show that these modified measures can be seen as an aggregation of measures of poverty and affluence.

There are only a few papers on multivariate polarization. Gigliarano and Mosler (2009b) generalize the idea of the ”decline of middle class” to higher dimensions. They calculate the volume of a middle class that contains a fixed portion of the population. An increasing volume of the middle class compared to the volume of the entire population is interpreted as a “decline of middle class”.

Our new measures of univariate polarization can be extended to a new definition of multivariate polarization, i.e. aggregating the distances of the people at the margins of the society to a given middle class region. Our index increases if a person leaves the middle class, if a person outside the middle class increases his or her distance to middle class and if people outside the middle class become homogenous.

This paper is organized as follows: Sections two and three shortly review the polarization indices of Wang and Tsui (2000) and some indices of poverty and affluence. Section four introduces the new class of univariate indices, a polarization ordering and also multivariate indices. Then, in Section five, the approach is applied to German data on income and working hours. Section six concludes.

[2]E.g. Blackburn and Bloom (1985) already defined a middle class interval for families, i.e. 60 to 225 percent of the median of families income (lower middle class: 60 to 100 percent, middle class 100 to 160 percent and upper middle class 160 to 225 percent).

2 Indices of Wang and Tsui

Consider a population of n individuals. Let

$$\mathcal{D} = \{\boldsymbol{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n_+ \setminus \{\mathbf{0}\} | 0 < x_1 \leq x_2 \leq \ldots \leq x_n\}$$

be the set of all ordered income distributions. To define the indices, we require the median income, denoted by $m(\boldsymbol{x})$.

Wang and Tsui (2000) define polarization measures by

$$\Psi_a(\boldsymbol{x}) := \frac{1}{n} \sum_{i=1}^{n} \psi\left(|x_i - m(\boldsymbol{x})|\right) \tag{1}$$

and

$$\Psi_r(\boldsymbol{x}) := \frac{1}{n} \sum_{i=1}^{n} \psi\left(\left|\frac{x_i - m(\boldsymbol{x})}{m(\boldsymbol{x})}\right|\right), \tag{2}$$

where ψ is a continuous function on $[0, \infty)$.

These indices are based on the distances (absolute or relative) to the median income $m(\boldsymbol{x})$. The distances are “weighted” by transforming them with a continuous function ψ. Finally, the mean of the weighted distances is calculated.

The two indices of Wang and Tsui are characterized by three important postulates of polarization, i.e. increasing spread, increasing bipolarity and “zero for equal incomes”.

First of all, the decline of middle class is a result of increasing distances from the median income, i.e. the central middle class income. Therefore an index that measures the decline of the middle class should increase if such changes in the income distribution occur:

Postulate 1 (Increasing spread (IS)) *The polarization index rises if an income above the median income increases or an income below the median income decreases.*

In inequality measurement a progressive transfer should increase the index. In line with this, a progressive transfer between a person with income below median income and another with income above median income decreases polarization because of postulate (IS). But this is not always the case: If some incomes above (below) the median become more equal by progressive transfers, i.e. they are more polarized, then the polarization index should increase. This is the main difference between polarization measurement and inequality measurement.

Postulate 2 (Increasing bipolarity (IB)) *The polarization index rises by a progressive transfer between two individuals both receiving incomes above the median income or between two individuals both receiving incomes below the median.*

The third postulate is the trivial condition that the index is zero if there is no inequality in income distribution, i.e. all people gain the same income:

Postulate 3 (Zero for equal incomes (Z)) *The polarization index is zero if all people receive the same income.*

Wang and Tsui (2000) showed the following characterization:[3]

Proposition 1 *Indices of the form Ψ_a and Ψ_r satisfy postulate (IS), (IB) and (Z) if and only if the ψ is a strictly increasing and strictly concave function on $\mathbb{R}_+$ and $\psi(0) = 0$.*

Figure 1 gives an example for measuring the contribution of each person to the polarization index, i.e. $f(y) = \psi\left(\left|\frac{y-m(\boldsymbol{x})}{m(\boldsymbol{x})}\right|\right)$, where ψ is strictly increasing and strictly concave with $\psi(0) = 0$, in case of relative polarization index and analogously for absolute measurement.

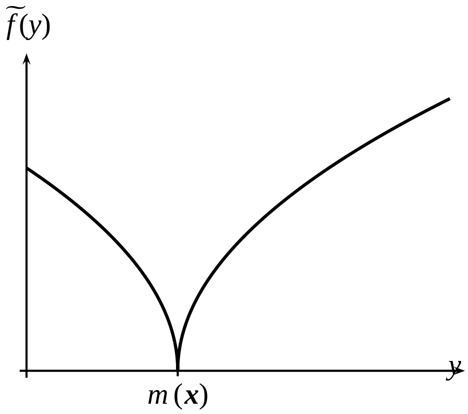

Figure 1: Example for $f(y)$: $\widetilde{f}(y) = \left|\frac{y-m(\boldsymbol{x})}{m(\boldsymbol{x})}\right|^{0.5}$.

[3]See Wang and Tsui (2000), Proposition 5.

3 Indices of poverty and affluence

In Section four we define polarization indices by replacing the median in the Wang and Tsui measures with a middle class interval. It will turn out that these new polarization indices are the sum of a poverty index and an affluence index. Therefore, we shortly take a look on the measurement of poverty and affluence in this section.

To construct a general class of income poverty indices in the usual way, we first have to identify the poor. This is done by a poverty line π, often a special percentage of the median income. Secondly, we have to quantify the degree of poverty of the poor by an increasing function of relative lack of income $\frac{\pi - x_i}{\pi}$. In this paper we use the class $\mathcal{P}$ of relative poverty indices, with elements

$$P(\boldsymbol{x}) = \frac{1}{n} \sum_{i=1}^{n} \psi_{poor} \left(\left(\frac{\pi - x_i}{\pi} \right)_+ \right),$$

where $\psi_{poor} : [0,1] \to \mathbb{R}_+$ is a (weakly) increasing function, say individual illfare function, and $(a)_+ := \max\{a, 0\}$. Note that many poverty indices like those of Foster et al. (1984), $P_{FGT}(\boldsymbol{x}) = \frac{1}{n} \sum_i \left(\left(\frac{\pi - x_i}{\pi} \right)_+ \right)^\alpha$, $\alpha > 0$, are included in this class $\mathcal{P}$. For a recent survey on poverty measurement see Chakravarty and Muliere (2004).

The measurement of affluence is less often investigated. Affluence indices can be constructed analogously to those of poverty: Firstly, identify the rich, i.e. all people with income above an affluence line ρ. Then, measure the extent of their richness. In Peichl, Schäfer and Scheicher (2010) such a class of relative affluence indices has been defined. Similar to this class, let

$$R(\boldsymbol{x}) = \frac{1}{n} \sum_{i=1}^{n} \psi_{rich} \left(\left(\frac{x_i - \rho}{\rho} \right)_+ \right),$$

where $\psi_{rich} : [0, \infty) \to \mathbb{R}_+$ is a (weakly) increasing function, say individual affluence function. E.g. for $\psi_{rich}(y) = \left(1 - \frac{1}{y+1}\right)^\alpha$, $\alpha > 0$, we obtain

$$R_\alpha(\boldsymbol{x}) = \frac{1}{n} \sum_i \left(\left(\frac{x_i - \rho}{x_i} \right)_+ \right)^\alpha. \quad (3)$$

Analogously, we can define absolute poverty and affluence indices.

4 New indices of polarization and polarization orderings

Now we define new indices of univariate and multivariate polarization and univariate polarization orderings.

4.1 New univariate indices of polarization

In our opinion polarization indices (1) and (2) have two drawbacks:

The first drawback is that variations of income near the median income, i.e. changes in the middle class, have a higher influence on the index than variations which are further away from $m(\boldsymbol{x})$. The reason for that delivers Postulate 2, the function ψ has to be strictly concave and strictly increasing. To illustrate this, we take a look at the following two income distributions:

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9
distribution A	0	0	0	100	100	100	200	200	200
distribution B	0	0	30	70	100	130	170	200	200

We obtain distribution B from A by two progressive transfers between x_3 and x_4 and between x_6 and x_7. Due to increasing bipolarity (Postulate 2) polarization increases from distribution A to B. But many people will rather claim highest polarization in distribution A, with totally homogenous lower, middle and upper classes. To get rid of this problem, we propose polarization measures that do not take the middle class, say the interval $[\pi, \rho]$, $\pi < \rho$, into account. Now the distribution of income inside the middle class will not affect the polariziation measure anymore. With this reduction of information we are able to focus on size and distribution of the outsiders (the poor and the rich). The proportion of the outsiders gives us directly information about the size of the middle class. Moreover the the distribution of the outsiders may be more or less homogenous. This allows us to evaluate their importance in democratic society.

The second problem is the skewness of income distribution, where incomes are bounded below by zero but do not have an upper bound. The summands in the Wang and Tsui index are bounded by $\psi\left(\left|\frac{0-m(\boldsymbol{x})}{m(\boldsymbol{x})}\right|\right) = \psi(1)$ for incomes below median income, but unbounded for incomes above median income. Therefore we propose to use two functions measuring the distance of the poor people to middle class and the distance of the rich people to the middle class, differently.

In contrast to Figure 1, Figure 2 shows an example how the share of each person to the polarization index could be measured after the two modifications.

Firstly, we define the modified absolute index

$$\Pi_a(\boldsymbol{x}) := \frac{1}{n}\left(\sum_i \psi_1\left((\pi - x_i)_+\right) + \sum_i \psi_2\left((x_i - \rho)_+\right)\right), \qquad (4)$$

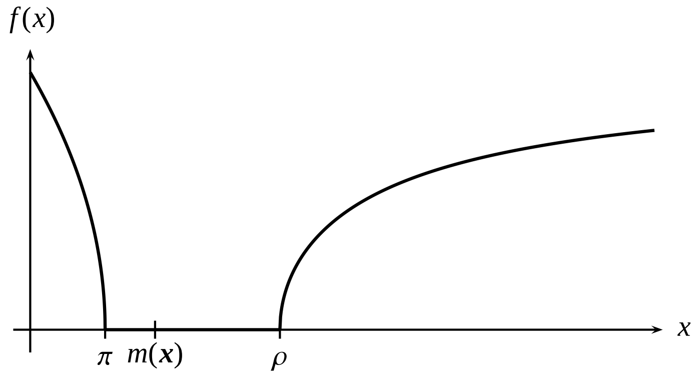

Figure 2: Modified function of Figure 1.

where ψ_1 and ψ_2 are weakly increasing functions. This index can be nicely interpreted as an aggregation of an absolute poverty index and an absolute affluence index, where ψ_1 and ψ_2 equal ψ_{poor} and ψ_{rich}. The simplest example is the "inequality" index of Stark (1972), which is a head count measuring the percentage of people who are either poor or rich:

$$\Pi_{Stark}(\boldsymbol{x}) = \frac{1}{n}\left(\sum_i \mathbf{1}_{\pi - x_i > 0} + \sum_i \mathbf{1}_{x_i - \rho > 0}\right),$$

where $\mathbf{1}_{\pi - x > 0} = \begin{cases} 1 & \text{if } \pi - x > 0, \\ 0 & \text{otherwise.} \end{cases}$

Secondly, a modified relative index is defined by

$$\Pi_r(\boldsymbol{x}) := \frac{1}{n}\left(\sum_i \psi_1\left(\left(\frac{\pi - x_i}{m(\boldsymbol{x})}\right)_+\right) + \sum_i \psi_2\left(\left(\frac{x_i - \rho}{m(\boldsymbol{x})}\right)_+\right)\right), \tag{5}$$

where ψ_1, ψ_2 are weakly increasing functions. Index (5) is also an aggregation of relative (to median income) indices of poverty and affluence.[4]

Two important questions have to be taken in consideration now: Firstly, what postulates should be satisfied for the new indices? And secondly, if the usual poverty and affluence indices are appropriate in respect to the postulates of polarization, or not.

[4]Note that the denominators of the poverty and affluence indices in Section 3 are different. This difference is not a problem, since mostly $\pi := c \cdot m(\boldsymbol{x})$, $c \in (0, 1)$ and $\rho := d \cdot m(\boldsymbol{x})$, $d > 1$ and the constants c and d can be included by choosing the function ψ_1 and ψ_2, respectively.

The new indices focus on incomes outside the interval $[\pi, \rho]$, therefore we have to modify the postulates 1-3:

Postulate 1* (Increasing spread outside $[\pi, \rho]$) *The polarization index rises if an income above the richness line ρ increases or an income below the poverty line π decreases.*

Postulate 2* (Increasing bipolarity outside $[\pi, \rho]$) *The polarization index rises by a progressive transfer between two individuals if both persons receive incomes above richness line or if both individuals receive incomes below poverty line.*

Postulate 3* (Zero for incomes in $[\pi, \rho]$) *The polarization index is zero if all people receive middle class incomes in $[\pi, \rho]$.*

Since we can interpret the new polarization indices as aggregation of poverty and affluence indices, many postulates of poverty measurement can be adopted to polarization measurement.

We review the standard technical postulates. They seem to be appropriate for most types of social indicators:

Postulate 4 (Anonymity) *The polarization index remains unchanged under a permutation of incomes.*

Postulate 5 (Replication invariance) *If the population is replicated several times, the polarization index will not change.*

Postulate 6 (Continuity) *The polarization index is continuous in the income vector.*

Now we adapt some normative postulates. In poverty (affluence) measurement one always focuses on poor (rich) people. This also seems to be useful in polarization, when a middle class interval is used instead of a middle class point. Only incomes outside the middle class interval should contribute to the polarization index:

Postulate 7 (Focus) *Given the middle class, the polarization index depends only on people with income outside the middle class.*

If there are no modifications to the income distribution, but (by some normative reasons) the poverty (affluence) line increases (decreases), then poverty (affluence) should increase. This should be the same in polarization measurement. A decreasing middle class interval increases the distance of each poor and each rich to the middle class. Moreover some people leave the middle class and become either poor or rich.

Postulate 8 (Decreasing middle class interval) *The polarization index increases if, ceteris paribus, the middle class interval decreases, i.e. the new middle class interval is a subset of the old one.*

Finally, polarization should change if an additional person enters the population. If this person earns a middle class income, then the portion of people outside the middle class decreases. Polarization should decrease. Or, if this is a non-middle class person, polarization should increase:

Postulate 9 (Poverty and richness growth) *The polarization index rises if a poor or a rich person enters the population.*

Postulate 10 (Middle class growth) *The polarization index decreases if a person that belongs to the middle class enters the population.*

The following theorem relates the definitions of polarization indices from above to some of these postulates.

Proposition 2 *Every polarization index of the form (4) and (5) satisfies the postulates of*

i) ***replication invariance*** *and* ***anonymity****,*

ii) ***zero for incomes in*** $[\pi, \rho]$ *and* ***focus,*** *if* $\psi_1(0) = 0$ *and* $\psi_2(0) = 0$*,*

iii) ***continuity*** *if* ψ_1 *and* ψ_2 *are continuous,*

iv) ***decreasing middle class interval*** *and* ***increasing spread outside*** $[\pi, \rho]$ *if* ψ_1 *and* ψ_2 *are non-negative and strictly increasing,*

v) ***increasing bipolarity outside*** $[\pi, \rho]$ *if* ψ_1 *and* ψ_2 *is strictly increasing and strictly concave.*

The second question, whether we should use the usual poverty and richness indices or different kinds of indices, has already been answered by Theorem 2. The indices of the form (5) satisfy the important postulates 1* and 2* if ψ_{poor} and ψ_{rich} are strictly increasing and strictly concave functions. E.g. the illfare functions of the FGT indices for $\alpha \in (0,1)$ or the individual affluence function of the richness index (3) for $\alpha \in (0,1)$. But these indices are not the sort of indices usually employed in poverty measurement, since a progressive transfer between two poor people should decrease, not increase the poverty index. This shows a clear difference between the usual measures of poverty and our indices of polarization. In Peichl et al. (2010) we discuss whether affluence indices should have concave or convex functions ψ_{rich}.

4.2 New polarization ordering

For the new indices of polarization we have to define a middle class interval. This is a disadvantage of our approach, but it also occurs in poverty measurement. The remedy for this drawback of poverty measurement is a poverty ordering that is uniform in poverty line, see Foster and Shorrocks (1988a,b).

Analogously, we define:

Definition 1 (Ordering uniform in middle class interval) *Consider poverty lines* $\pi \in [\pi_{min}, \pi_{max}]$ *and affluence lines* $\rho \in [\rho_{min}, \rho_{max}]$, *with* $\pi_{max} \leq \rho_{min}$, *and an index* Π. *For* $\boldsymbol{x}$ *and* $\boldsymbol{y} \in \mathcal{D}$ *define the polarization ordering*

$$\boldsymbol{y} \prec_{\mathcal{M}} \boldsymbol{x}$$

if $\Pi(\boldsymbol{y}) \leq \Pi(\boldsymbol{x})$ *holds for all middle class intervals* $[\pi, \rho] \in \mathcal{M}$, *with* $\mathcal{M} = \{[\pi,\rho] \,|\, [\pi_{max}, \rho_{min}] \subset [\pi, \rho] \subset [\pi_{min}, \rho_{max}]\}$.

More applicable is the following weaker definition, where we restrict ourselves to a finite collection of nested middle class intervals.

Definition 2 (Ordering uniform for a nest of middle class intervals) *Consider a sequence of nested middle intervals* $[\pi_1, \rho_1] \supset [\pi_2, \rho_2] \supset ... \supset [\pi_n, \rho_n]$ *and an index* Π. *For* $\boldsymbol{x}$ *and* $\boldsymbol{y} \in \mathcal{D}$ *define the polarization ordering*

$$\boldsymbol{y} \prec_{[\pi_i, \rho_i]} \boldsymbol{x}$$

if $\Pi(\boldsymbol{y}) \leq \Pi(\boldsymbol{x})$ *holds for all middle class intervals* $[\pi_i, \rho_i]$, $i = 1, 2, \ldots, n$.

In Section 5 we illustrate the ordering $\prec_{[\pi_i, \rho_i]}$ with german data.

4.3 New multivariate indices of polarization

Sometimes it makes sense to consider more than one attribute, e.g. students are mostly poor in income but rich in education. Do they belong to middle class? Probably many people would say so. So far there are only a few papers dealing with multivariate polarization, see Gigliarano and Mosler (2009a,b).

Let $\boldsymbol{X} = \left((x_{11}, \ldots x_{1k})^T, \ldots (x_{n,1}, \ldots x_{nk})^T\right) \in \mathbb{R}_+^{k \times n}$ be the distribution of a population with k attributes and n individuals.

The simplest way to consider more than one attribute is to calculate a vector of univariate polarization indices. But then the above mentioned student contributes to both univariate polarization indices. S/he is poor in income and rich in education and therefore s/he does not belong to the income middle class and the education middle class. But it may be sensible that s/he does not contribute to the polarization index, because s/he will belong to the income middle class in future due to education.

To cope with this problem we define multivariate indices in two steps: First, we have to define middle class region $\mathcal{M}$, see figure 3:

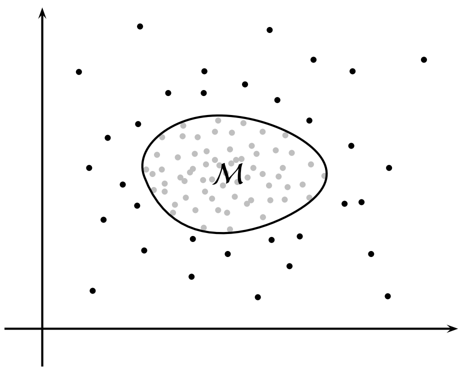

Figure 3: A middle class region $\mathcal{M}$.

Some observation are outside the middle class. Therefore we have to evaluate the situation of those people, in a second step.

This is done by calculating a somehow defined distance $d(\boldsymbol{x}_i, \mathcal{M})$ between the individuum i and the middle class $\mathcal{M}$. Now we are able to define multivariate polarization indices analog to our univariate polarization indices (4):

$$\Psi(\boldsymbol{X}) = \frac{1}{n} \sum_i f(d(\boldsymbol{x}_i, \mathcal{M})), \tag{6}$$

where $f : \mathbb{R}_+ \to \mathbb{R}_+$ is a continuous, increasing and concave function, with $f(0) = 0$. Since the distance to the middle class in general is not bounded, we could use increasing and concave function f that maps on $[0, 1]$ to guarantees a normalized index Ψ.[5]

The easiest example of a multivariate polarization index Ψ is the proportion of people not belonging to the middle class $\mathcal{M}$, the head count:

$$\Psi_{HC}(\boldsymbol{X}) = \frac{1}{n} \sum_i \mathbf{1}_{(x_{i1},\ldots,x_{ik})^T \notin \mathcal{M}} \tag{7}$$

Analogue to the univariate case, we can define multivariate polarization orderings for a nested sequence of middle class regions.

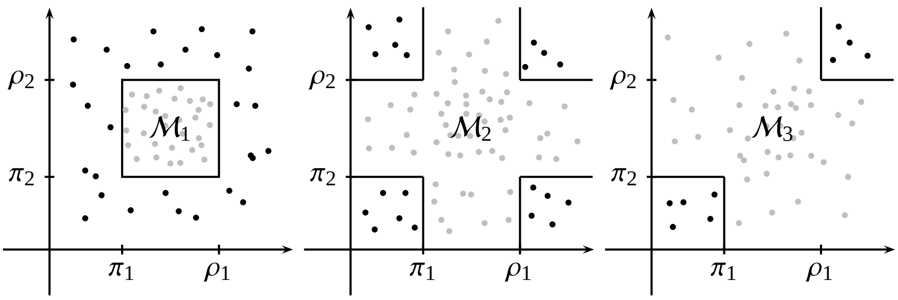

Figure 4: Different middle class regions $\mathcal{M}_1$,$\mathcal{M}_2$, $\mathcal{M}_3$.

We illustrate our multivariate indices with some examples: Figure 4 shows three simple examples of middle class regions $\mathcal{M}_i$, $i = 1, 2, 3$. Middle class $\mathcal{M}_1$ includes all people that belong to all univariate middle classes, i.e. $[\pi_1, \rho_1]$ and $[\pi_2, \rho_2]$ respectively. $\mathcal{M}_2$ includes all people that belongs to at least one univariate middle class. Finally, all people belong to the middle class $\mathcal{M}_3$, if they are neither poor in both attributes nor rich in both attributes. Of course we can imagine many other middle class regions, but we restrict ourselves to this middle classes, since the following calculations will be complicated otherwise.

The choice of the middle class region and multivariate index depends on the fact whether the attributes are substitutable or complementary. E.g. if the attributes are substitutable, then a low value in one attribute is compensated by a high value of another attribute.[6]

[5]See our detailed discussion in Peichl et al. (2010), whether distances to the richness line should be evaluated with convex or concave functions.

[6]For multivariate poverty measurement see e.g. the papers of Bourguignon and Chakravarty (2003) and Tsui (2002).

To evaluate the situation of non middle class people in those different situations, let

$$d(x,[a,b]) = \begin{cases} \min\{|x-a|,|x-b|\} & \text{if } x \notin [a,b], \\ 0 & \text{if } x \in [a,b] \end{cases}$$

be the distance between a number x and an interval $[a,b]$. To calculate the distance between $\boldsymbol{x}_i = (x_{i1},\ldots,x_{ik})$ and $\mathcal{M}_1$ we use the L_1 norm:

$$d(\boldsymbol{x}_i,\mathcal{M}_1) = \sum_{j=1}^{k} d(x_{ij},[\pi_j,\rho_j]) \,.$$

For $\mathcal{M}_2$ we calculate

$$d(\boldsymbol{x}_i,\mathcal{M}_2) = \min_j(d(x_{ij},[\pi_j,\rho_j])) \,,$$

and for $\mathcal{M}_3$:

$$d(\boldsymbol{x}_i,\mathcal{M}_3) = \begin{cases} \min_j(\pi_j - x_{ij}) & \text{if } \boldsymbol{x}_i \in (-\infty,\pi_1)\times(-\infty,\pi_2)\,, \\ \min_j(x_{ij}-\rho_j) & \text{if } \boldsymbol{x}_i \in (\rho_1,\infty)\times(\rho_2,\infty)\,, \\ 0 & \text{otherwise.} \end{cases}$$

The next section will illustrate the new indices and ordering.

5 Polarization in Germany

An advantage of the new indices is, that they allow us to understand polarization as a combination of poverty and richness measures. With the GSOEP (German Socio-Economic Panel) data we analyse now whether the new indices perform at least as good as the common measures of declining middle class, i.e. the Wang and Tsui (2000) index.

5.1 Data

The GSOEP is a panel study of private households in Germany since 1984. We use the "cross-national equivalent files" (CNEF) of GSOEP.[7] In these files for each year a variable "household post-government income" (that refers to the previous

[7]The data used in this publication are provided by the Cross-National Equivalent File (CNEF) project at the College of Human Ecology at Cornell University, Ithaca, N.Y.

year incomes) is already calculated. We divide this household income by a factor that is calculated with a modified OECD equivalence scale, i.e. 1.0 for household head plus 0.3 for each child younger than 15 years plus 0.5 for each other household member, to obtain a personal equivalence income.

For more information on the GSOEP see Haisken De-New and Frick (2005) and for further information on the CNEF see Lillard et al. (2008).

5.2 Univariate polarization indices

Firstly, we calculate the relative Wang and Tsui index $\Psi_r(\boldsymbol{x})$ with $\psi(y) = y^{0.5}$. We decompose it in

$$\Psi_r(\boldsymbol{x}) = \Psi_1(\boldsymbol{x}) + \Psi_2(\boldsymbol{x}) + \Psi_3(\boldsymbol{x}),$$

where $\Psi_1(\boldsymbol{x}) := \frac{1}{n}\sum_{i:x_i\in[\pi,\rho]}\psi\left(\left|\frac{x_i-m(\boldsymbol{x})}{m(\boldsymbol{x})}\right|\right)$ is the part of polarization that comes from the middle class, $\Psi_2(\boldsymbol{x}) := \frac{1}{n}\sum_{i:x_i<\pi}\psi\left(\left|\frac{x_i-m(\boldsymbol{x})}{m(\boldsymbol{x})}\right|\right)$ from the poor and $\Psi_3(\boldsymbol{x}) := \frac{1}{n}\sum_{i:x_i>\rho}\psi\left(\left|\frac{x_i-m(\boldsymbol{x})}{m(\boldsymbol{x})}\right|\right)$ from the rich.

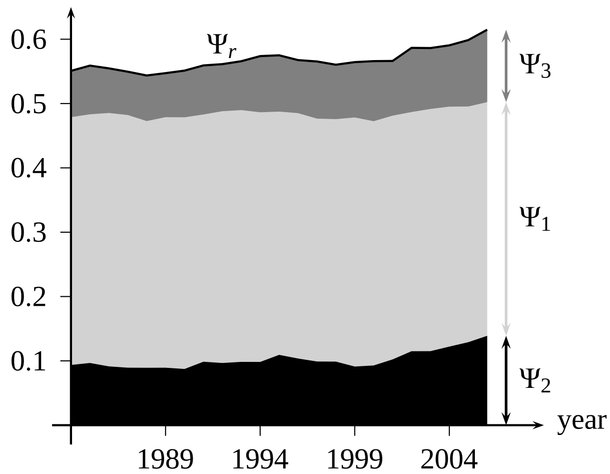

Figure 5: Ψ_r, with $\psi(y) = y^{0.5}$ for Germany, with contributions of middle class (Ψ_1), poor (Ψ_2) and rich (Ψ_3).

Figure 5 shows that most of the polarization (between 0.54 and 0.61) is obtained from the incomes in the middle class (between 0.36 and 0.39) and this amount is nearly constant, therefore we do not take this amount into account. A smaller amount that includes most of the variation comes from the incomes outside the middle class (0.16 to 0.25).[8]

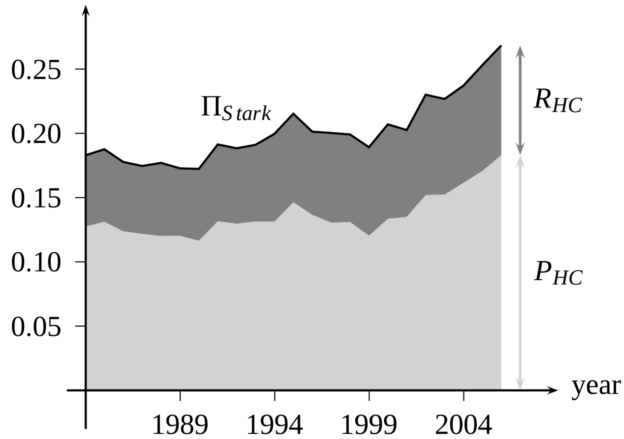

Figure 6: Proportion of poor (P_{HC}), rich (R_{HC}) and the Stark index (Π_{Stark}) for Germany.

We find this enormous change in income distribution already if we look at the percentage of people outside the middle class, i.e. the Stark index.[9] This index increases from 20% in 2001 to 27% in 2006, since for the same years the percentage of poor increases from 14% to 18% and the percentage of rich from 7% to 9%, see Figure 6 (and Table 4 in the Appendix).

We observe similar results for another index of our class of new polarization indices,

$$\Pi_{\alpha_1,\alpha_2}(\boldsymbol{x}) := \underbrace{\frac{1}{n}\sum_i \left(\left(\frac{\pi - x_i}{\pi}\right)_+\right)^{\alpha_1}}_{=:\Pi^{poor}(\boldsymbol{x})} + \underbrace{\frac{1}{n}\sum_i \left(\left(\frac{x_i - \rho}{x_i}\right)_+\right)^{\alpha_2}}_{=:\Pi^{rich}(\boldsymbol{x})},$$

with $\alpha_1, \alpha_2 \in (0,1)$.[10] See Figure 7 (and Table 5 in the Appendix) for the results for $\alpha_1 = \alpha_2 = 0.5$.

The next step is to investigate how changes in the income distribution will be represented by different choices of individual illfare and affluence function, e.g. for

[8] See Table 3 in the Appendix for the numerical results.

[9] Note that the "inequality" index of Stark (1972) is already one of the new kind of polarisation indices.

[10] We use the same poverty and richness lines as in Peichl et al. (2010), i.e. 60% and 200% of median income.

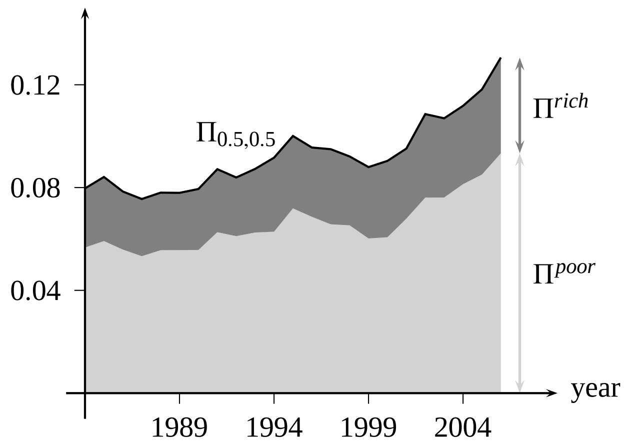

Figure 7: New index $\Pi_{0.5,0.5}$ for Germany.

different α_1 and α_2 in $\Pi^{poor}_{\alpha_1}$ and $\Pi^{rich}_{\alpha_2}$, respectively; see Table 1. We find smallest relative variation for $\Pi^{poor}_{\alpha_1=0.1}$ (from 0.103 in 1986 to 0.157 in 2006, i.e. 52.4 % rise) and largest relative variation for $\Pi^{poor}_{\alpha_1=0.9}$ (from 0.035 in 1986 to 0.061 in 2006, i.e. 74.3 % rise). For $\Pi^{rich}_{\alpha_2}$ with $\alpha_2 = 0.1$, 0.5 and 0.9 we find a 62.0, 60.9 and 69.2 % rise, respectively. If we use these poverty and affluence measures to construct polarization measures, we find the smallest relative rise for $\Pi_{0.1,0.5}$ (from 0.126 in 1986 to 0.194 in 2006, i.e. 54.0 % rise) and for $\Pi_{0.1,0.1}$ (from 0.147 in 1986 to 0.228 in 2006, i.e. 55.1 % rise). The largest rise is obtained for $\Pi_{0.9,0.5}$ (69.0 %) and $\Pi_{0.9,0.9}$ (75.0 %). Therefore we are able to construct indices that take poverty or affluence more or less into account. This depends on ethical judgments whether the polarization that occurs due to the poor or due to the rich is more problematic for society.

5.3 Polarization Ordering

We also defined a new polarization ordering $\prec_{[\pi_i,\rho_i]}$: Now we calculate polarization $\Pi_{\alpha_1=0.5,\alpha_2=0.5}$ for the years 1986, 1996 and 2006 for middle class interval $[\pi_i, \rho_i]$ with $\pi_i = 0.4 + 0.01 \cdot i$ and $\rho_i = 2.4 - 0.02 \cdot i$ times the median income, $i = 0, 1, 2, \ldots, 40$, i.e. $[0.4, 2.4] \supset [0.41, 2.38] \supset \ldots \supset [0.8, 1.6]$. Figure 8 shows that $\boldsymbol{x}^{1986} \prec_{[\pi_i,\rho_i]} \boldsymbol{x}^{1996} \prec_{[\pi_i,\rho_i]} \boldsymbol{x}^{2006}$. We detect that that for each of these middle class intervals polarization is higher in 1996 than in 1986. This is of course not surprising, since Germany was reunited in 1990. More interesting is the enormous increase of polarization from 1996 to 2006.

		$\alpha_1, \alpha_2 = .1$	$\alpha_1, \alpha_2 = .5$	$\alpha_1, \alpha_2 = .9$
1986	$\Pi_{\alpha_1}^{poor}$	0.103	0.056	0.035
	$\Pi_{\alpha_2}^{rich}$	0.044	0.023	0.013
	Π_{α_1,α_2}	0.147	0.78	0.048
1996	$\Pi_{\alpha_1}^{poor}$	0.116	0.069	0.046
	$\Pi_{\alpha_2}^{rich}$	0.053	0.027	0.015
	Π_{α_1,α_2}	0.170	0.096	0.061
2006	$\Pi_{\alpha_1}^{poor}$	0.157	0.093	0.061
	$\Pi_{\alpha_2}^{rich}$	0.071	0.037	0.022
	Π_{α_1,α_2}	0.228	0.131	0.084

Table 1: Π_{α_1,α_2} for various α_1, α_2.

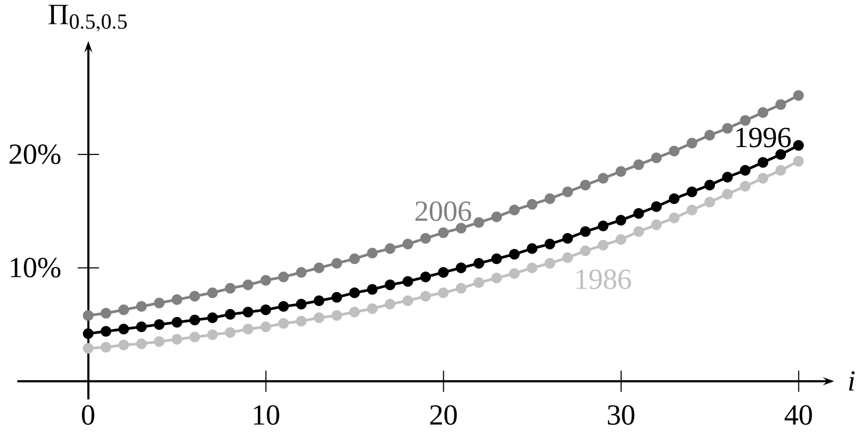

Figure 8: Polarization ordering $\prec_{[\pi_i,\rho_i]}$, with $\pi_i = 0.4 + 0.01 \cdot i$ and $\rho_i = 2.4 - 0.02 \cdot i$ times the median income, $i = 0, 1, 2, \ldots, 40$.

5.4 Multivariate polarization indices

In this section we illustrate the new multivariate polarization indices by the following example. Consider a person with an average income. Does this person really belong to the middle class? If s/he works an average time to gain an average income we would say so. But maybe s/he works only a few hours per month, or – probably more realistic – s/he works many hours overtime to reach an average income. Unfortunately in each case the person would belong to middle class in univariate measurement (regarding income).

Therefore, as a multivariate application we investigate employed (part time and full time) single person households only[11] and use the variables "household post-government income" and "annual work hours of individual"[12]. As mentioned above, we do not only focus on income and take the working hours into account, we define the middle class by all people who work an average time **and** gain an average income. This is done because of the following normative specifications: A person has to gain an average income to belong to the middle class. Moreover s/he will not be included in the middle class, if s/he either has to work overtime or to work only a few hours to receive an average income, because these people will either belong to the poor or the rich, if they do average hours of work. Therefore we use middle class $\mathcal{M}_1$.[13] The multivariate polarization measure (6) for $f(y) = y^{0.5}$ is:

$$\Psi_{\mathcal{M}_1}(\boldsymbol{X}) = \frac{1}{n}\sum_i (d(\boldsymbol{x}_i, \mathcal{M}_1))^{0.5} \ .$$

Let poverty and affluence lines be 60% and 200% of median income and 75% and 125% of median hours of work. Figure 9 shows the middle class region and a 25% sample of data of 2006.

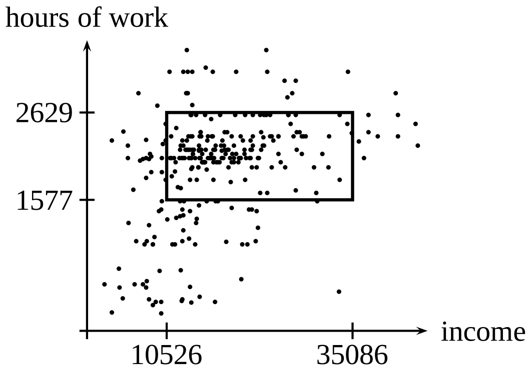

Figure 9: Middle class region and 25% sample of data of 2006.

Table 2 shows the results for several years. Obviously the univariate polarization indices sometimes show the same development (compare 1996 and 2001) and sometimes they do not (compare changes of Π^{income}_{Stark} and $\Pi^{workhours}_{Stark}$ between 1986 and 1991, or of $\Pi^{income}_{0.5,0.5}$ and $\Pi^{workhours}_{0.5,0.5}$ between 1986 and 2006).

[11]Note, this example is only an illustration for multivariate indices, because there are only few observations available (e.g. 582 in 1986).

[12]For more information on this variables see Lillard et al. (2008).

[13]Note that different normative specifications would lead to other middle class regions.

	1986	1991	1996	2001	2006
median income	15566	17684	16617	16709	17543
median work hours	2078	2078	2078	2078	2103
Π^{income}_{Stark}	19.6	20.2	20.5	23.7	23.2
$\Pi^{work\,hours}_{Stark}$	34.1	32.8	36.0	39.7	30.0
Ψ_{HC}	41.6	40.9	43.5	47.1	42.7
$\Pi^{income}_{0.5,0.5}$	10.0	9.3	9.7	11.9	11.4
$\Pi^{work\,hours}_{0.5,0.5}$	16.5	14.6	15.8	20.0	15.4
$\Psi_{\mathcal{M}_1}$	21.5	21.5	22.2	27.2	24.7

Table 2: Multivariate (Ψ_{HC}, $\Psi_{\mathcal{M}_1}$) and univariate indices (in %).

6 Concluding remarks

In this paper we defined new polarization indices that satisfy several important postulates. In the univariate case, they modify the indices of Wang and Tsui (2000) and in the multivariate case they provide an alternative to the approach of Gigliarano and Mosler (2009b), as we do not calculate the volume of the middle class region, but the distances to the middle class region.

Now we are able to define polarization by poverty and affluence. Although the poverty and affluence indices do not correspond to the usual type, the commonplace that “society is more polarized, since the rich get richer and the poor get poorer,” can be interpreted much better.

A result of the empirical illustration is an increase of polarization in Germany, especially in the years since 2000. The percentage of poor and rich people (the Stark index) increases from 18% in 1984 to 27% in 2006, i.e. a surplus of 50%. This enormous change in the income distribution is not fully visible by a Wang and Tsui type index, since people with income near the median contribute heavily to the index. Here the Wang and Tsui type index calculates only changes from 0.55 in 1984 to 0.61 in 2006. This is just a surplus of 11%. The new index shows this change in income distribution more clearly; it increases from 0.08 in 1984 to 0.13 in 2006, a 64% surplus.

Different ethical judgments about changes in poverty and changes in affluence can be integrated into the measurement now, since the new indices give us the opportunity to take changes in poor person´s income and rich person´s income more or less into account.

Polarization orderings allows us to state that polarization increases from 1986 to 1996 and even more from 1996 to 2006 for several definitions of middle class interval.

The illustration of a multivariate polarization index shows that the multivariate measurement may lead to different results in comparison to the univariate measures.

Especially for multivariate polarization further research is necessary: Firstly, a discussion on which postulates should be fulfilled, e.g. which kind of progressive transfer should increase multivariate polarization? And secondly, how can distances be calculated to more complex middle class sets?

7 Appendix

The following tables show the numerical results for the figures 5 to 7 in Section 5.

	84	85	86	87	88	89	90	91	92	93	94	95
$\Psi_{r,1}$	.39	.39	.39	.39	.38	.39	.39	.38	.39	.39	.39	.38
$\Psi_{r,2}$	.09	.10	.09	.09	.09	.09	.09	.10	.10	.10	.10	.11
$\Psi_{r,3}$	.07	.08	.07	.07	.07	.07	.07	.08	.07	.08	.09	.09
Ψ_r	.55	.56	.55	.55	.54	.55	.55	.56	.56	.57	.57	.58

	96	97	98	99	00	01	02	03	04	05	06
$\Psi_{r,1}$	.38	.38	.38	.39	.38	.38	.37	.38	.37	.37	.36
$\Psi_{r,2}$	.10	.10	.10	.09	.09	.10	.11	.12	.12	.13	.14
$\Psi_{r,3}$	.08	.09	.08	.09	.09	.08	.10	.09	.10	.10	.11
Ψ_r	.57	.57	.56	.56	.57	.57	.59	.59	.59	.60	.61

Table 3: Ψ_r, with $\psi(y) = y^{0.5}$ for Germany, with contributions of middle and non-middle class.

	84	85	86	87	88	89	90	91	92	93	94	95
P_{HC}	13	13	12	12	12	12	12	13	13	13	13	15
R_{HC}	6	6	5	5	6	5	6	6	6	6	7	7
Π_{Stark}	18	19	18	17	18	17	17	19	19	19	19	22

	96	97	98	99	00	01	02	03	04	05	06
P_{HC}	14	13	13	12	13	14	15	15	16	17	18
R_{HC}	6	7	7	7	7	7	8	7	8	8	9
Π_{Stark}	20	20	20	19	21	20	23	23	24	25	27

Table 4: Proportion of poor (P_{HC}), rich (R_{HC}) and the Stark index (Π_{Stark}) for Germany, in %.

	84	85	86	87	88	89	90	91	92	93	94	95
$\Pi_{0,5}^{poor}$	5.7	5.9	5.6	5.3	5.6	5.6	5.6	6.3	6.1	6.2	6.3	7.2
$\Pi_{0.5}^{rich}$	2.3	2.5	2.3	2.2	2.2	2.2	2.4	2.5	2.3	2.5	2.9	2.8
$\Pi_{0.5,0.5}$	8.0	8.4	7.8	7.6	7.8	7.8	7.9	8.7	8.4	8.7	9.2	10.0

	96	97	98	99	00	01	02	03	04	05	06
$\Pi_{0,5}^{poor}$	6.9	6.6	6.5	6.0	6.1	6.8	7.6	7.6	8.1	8.5	9.3
$\Pi_{0.5}^{rich}$	2.7	2.9	2.7	2.8	3.0	2.7	3.2	3.1	3.0	3.3	3.7
$\Pi_{0.5,0.5}$	9.6	9.5	9.2	8.8	9.0	9.5	10.9	10.7	11.2	11.8	13.1

Table 5: $\Pi_{\alpha_1=0.5,\alpha_2=0.5}$ for Germany, in %.

References

Blackburn, M. L. and Bloom, D. E. (1985). What is Happening to the Middle Class. *American Demographics* **7**, 18–25.

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

CHAKRAVARTY, S. R. and MAJUMDER, A. (2001). Inequality, polarization and welfare: Theory and applications. *Australian Economic Papers* **40**, 1–13.

CHAKRAVARTY, S. R. and MULIERE, P. (2004). Welfare indicators: a review and new perspectives. 2. Measurement of poverty. *Metron* **LXII**, 247–281.

D´AMBROSIO, C. (2001). Household characteristics and the distribution of income in Italy: An application of social distance measures. *Review of Income and Wealth* **47**, 43–64.

DUCLOS, J.-Y., ESTEBAN, J. and RAY, D. (2004). Polarization: Concepts, measurement, estimation. *Econometrica* **72**, 1737–1772.

ESTEBAN, J., GRADÍN, C. and RAY, D. (2007). An extension of a measure of polarization, with application to the income distribution of five OECD countries. *Journal of Economic Inequality* **5**, 1–19.

ESTEBAN, J.-M. and RAY, D. (1994). On the Measurement of Polarisation. *Econometrica* **62**, 819–851.

FOSTER, J., GREER, J. and THORBECKE, E. (1984). A Class of Decomposable Poverty Measures. *Econometrica* 761–766.

FOSTER, J. and SHORROCKS, A. (1988a). Poverty Orderings. *Econometrica* **56**, 173–177.

FOSTER, J. and SHORROCKS, A. F. (1988b). Poverty Orderings and Welfare Dominance. *Social Choice and Welfare* **5**, 179–198.

GIGLIARANO, C. and MOSLER, K. (2009a). Constructing indices of multivariate polarization. Journal of Economic Inequality.

GIGLIARANO, C. and MOSLER, K. (2009b). Measuring Middle-Class Decline in One and Many Attributes. Quaderno di Ricerca n. 333, Università Politecnica delle Marche.

HAISKEN DE-NEW, J. and FRICK, J., eds. (2005). *DTC - Desktop Compendium to The German Socio Economic Panel (SOEP)*. DIW, Berlin, www.diw.de, Version 8.0.

LILLARD, D. R., GRABKA, M. and FEIDIN, S. (2008). *Codebook for the Cross-National Equivalent File 1980-2005 BHPS – GSOEP – HILDA – PSID – SLID*. http://www.human.cornell.edu/che/PAM/Research/Centers-Programs/German-Panel/cnef.cfm, download April 08.

PEICHL, A., SCHAEFER, T. and SCHEICHER, C. (2010). Measuring Richness and Poverty – A micro data application to Europe and Germany. *forthcoming in Review of Income and wealth* .

STARK, T. (1972). *The Distribution of Personal Income in the United Kingdom 1949-1963*. Cambridge University Press, London.

TSUI, K.-Y. (2002). Multidimensional Poverty Indices. *Social Choice and Welfare* **19**, 69–93.

WANG, Y.-Q. and TSUI, K.-Y. (2000). Polarisation Orderings and New Classes of Polarization Indizes. *Journal of Public Economic Theory* **2**, 349–363.

WOLFSON, M. C. (1994). When Inequality Diverge. *The American Economic Review* **84**, 353–358.

WOLFSON, M. C. (1997). Divergent inequalities: Theory and empirical results. *Review of Income and Wealth* **43**, 401–421.