ECONSTOR

WWW.ECONSTOR.EU

Der Open-Access-Publikationsserver der ZBW – Leibniz-Informationszentrum Wirtschaft
The Open Access Publication Server of the ZBW – Leibniz Information Centre for Economics

Peichl, Andreas; Ochmann, Richard

Working Paper

Measuring distributional effects of fiscal reforms

ZBW Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Mitglied der
Leibniz-Gemeinschaft

Measuring Distributional Effects
of Fiscal Reforms

by

Richard Ochmann[*] and Andreas Peichl[‡]

September 2006

[*]   Berlin Doctoral Program in Economics and Management Science, Berlin, Germany.
richard.ochmann@gmx.de

[‡]Center for Public Economics, University of Cologne, Albertus-Magnus-Platz, 50923 Cologne, Germany.
a.peichl@uni-koeln.de

Abstract

The purpose of this paper is to provide an overview of how to analyse the distributional effects of fiscal reforms. Thereby, distributional effects shall be differentiated by four subconcepts, i.e. 1.) the traditional concept of inequality, 2.) the rather novel concept of polarisation, 3.) the concept of progression in taxation, and 4.) the concepts of income poverty and richness. The concept of inequality and the concept of income poverty are the by far most widely applied concepts in empirical analyses, probably since they appear to be the most transparent ones in their structure as well as the most controversial ones in political affairs. However, the concepts of richness, polarisation and progression in taxation shall additionally be subject of this analysis, since they appear to be useful devices on the course of analysing cause and effect of the other two concepts.

Contents

List of Tables

List of Figures

List of Acronyms

BMF Bundesministerium der Finanzen (German Federal Ministry of Finance)

CHU Clark, Hemming, Ulph

CV coefficient of variation

DAD Distributive Analysis/Analysing Distributive

DIW Deutsches Institut fuer Wirtschaftsforschung

EVS Einkommens- und Verbrauchsstichproben (incomes and expenditures surveys of the German Federal Statistical Office)

FALESt98 scientific use file factually anonymized individual income tax statistic

FGT Foster, Greer, Thorbecke

FiFo Finanzwissenschaftliches Forschungsinstitut an der Universitaet zu Koeln (Cologne Center for Public Economics)

GE generalized entropy

GDR German Democratic Republic

IR Income Redistributive

LESt individual income tax

LESt98 an individual-income-tax statistic sample of the German Federal Statistical Office

LVAR logarithmic variance

LIS Luxembourg Income Study

MLD mean logarithmic deviation

OECD Organization for Economic Cooperation and Development

RMD relative mean deviation

RR reranking

SDL standard deviation of the logarithms

SOEP Socio Economic Panel of the Deutsches Institut für Wirtschaftsforschung

SVR Sachverstaendigenrat zur Begutachtung der gesamtwirtschaftlichen Lage (an economic council of the German government)

TR Tax Redistributive

VAR variance

VARL variance of the logarithms

VE vertical equity

1 Introduction

Reforms of the troubled welfare state and especially the tax and benefit system are high on the political agenda in various countries. Such reforms usually affect the structure of tax burdens and the amount of benefits received and thus the distribution of disposable incomes. Since in a democratic political system, such reform proposals need to win the majority of votes before they can be implemented, it appears crucial to analyse who may gain and who may lose as a consequence of such reforms. This paper will therefore focus on the measurement of the distributional effects of policy reforms.

Empirical approaches use real data sets on income distributions in order to analyse distributional effects of various reform scenarios. Such approaches may be differentiated by the timing of their analysis. While ex-post approaches rather use actual data after the reform has been implemented, in order to evaluate its effects, ex-ante approaches simulate data with respect to an oncoming implementation of a reform and forecast potential effects of its implementation. The latter approach uses data provided by microsimulation models[1] to estimate the effects of future reforms with econometric methods.

The purpose of this paper is to provide an overview of how to analyse the distributional effects of fiscal reforms. Thereby, distributional effects shall be differentiated by four subconcepts, i.e. 1.) the traditional concept of inequality, 2.) the rather novel concept of polarisation, 3.) the concept of progression in taxation, and 4.) the concept of income poverty and complementarily richness. The traditional concept of inequality and the concept of income poverty are the by far most widely applied concepts in empirical analyses, probably since they appear to be the most transparent ones in their structure as well as the most controversial ones in political affairs. However, the concepts of richness, polarisation and progression in taxation shall additionally be subject of this analysis, since they appear to be useful devices on the course of analysing cause and effect of the other two concepts.

Firstly, it appears reasonable to limit an analysis of distributional effects to incomes, rather than applying it to total assets, primarily due to a better availability of data on incomes compared to data on total assets of people. Secondly, it appears necessary to define an appropriate concept of income to apply an analysis to. Following a concept of economic income, i.e. considering incomes as they have actually been generated on markets, yields a concept of *pre-government income*. Thereby, the sum of earnings generated from independent and dependent personal services, private assets as well as private transfers is called the *market income*.[2] Based on market incomes, post-government incomes in economic

---

[1]Microsimulation models on the quantification of distributional effects of tax reforms have firstly been applied by Orcutt (1957) and later on further developed by Orcutt et al. (1986). Cf. Gupta and Kapur (2000) for an introduction to the field of microsimulation and Peichl (2005) for an overview of the method of simulation to evaluate tax reforms.

[2]Employer contributions to compulsory health insurances, to compulsory long term care insurances, to unemployment insurances, and to the statutory pension insurance provisions, are thereby added to earnings from dependent personal services, since they are earned on markets as well and represent first

terms are derived by taking governmental payments into consideration. On the one hand income tax liabilities and social security contributions are deducted, and on the other hand pensions from the statutory pension insurance as well as social transfers[3] are added. The resulting difference between market incomes and *post-government incomes* may be interpreted as the result of governmental redistribution. However, it appears relevant to take into account that income units in tax statistics usually represent incomes of more than one person together, or they stem from a single person but are later on distributed among multiple members of a household[4], so that an analysis should allow for differences in the income units' needs, i.e. the population of income units is actually heterogeneous, consisting of singles, couples and families. In general, so called equivalence scales reflect both, economies of scale in household size, and differences in household characteristics, such as needs, location, age, number and age of children, and health. The most widely applied concepts of equivalence scales exhibit simple scale parameters.[5] E.g. the equivalence scales from the OECD attach weights to household members in relation to their age.[6]

Having defined the proper units of assessment as well as the appropriate concept of income, one is well equipped to analyse the distributional effects of fiscal reforms. In the course of this paper, several different concepts to measure the impact of such reforms on the income distribution are presented. The setup of the paper is organised as follows: Chapter 2 opens up the distributional analysis with inequality as the first one of the four distributional concepts to be analysed. Chapter 3 then follows with

---

governmental actions. Thus, the concept of market income denotes incomes prior to any governmental payments, i.e. to say in other words, the amount of gross incomes employees signed a contract for.

[3]Social transfers denote child benefits, child-rearing benefits, education benefits for students, unemployment compensation, housing benefits and social assistance benefits, and regular, but not irregular, supplementary grants.

[4]Tax units may denote incomes of couples in case they are assessed to the individual income tax by pair, or they may denote incomes of whole households, in case the other members of the household attached to the income recipient(s) of the tax unit considered do not gain any additional incomes on their own.

[5]The concepts of equivalence scales adapted by Buhmann et al. (1988) and Coulter et al. (1992) take into account economies of scale that occur in household needs as equity-relevant non-income differences between persons, in relation to disposable incomes of households. They distinguish household types by their size, with s, $s = 1,...n$ denoting the number of persons per household. Then, households are grouped by their size, so that p_s follows as the population share of group s, with $\sum_{s=1}^{n} p_s = 1$. Unadjusted, disposable incomes are continuously distributed over $[a, b]$ within each group, with density function $f_s(X)$ and distribution function $F_s(X)$. Coulter et al. (1992) define a simple equivalence scale rate for household s as $M_s = M(s, \theta)$, with $\theta \geq 0$ denoting the scale relativity parameter and M_s being an increasing function in s and in θ, with $M_1 = 1$. Thereby, scale relativities are defined in relation to the scale of a single-person household. The greater θ, ceteris paribus, the greater is the scale rate, i.e. the greater are the needs assumed for the multiple-person household relatively to a single-person household, independently of income. It then results $Y = \frac{X}{M_s}$ as the equivalent income for a person in group s, i.e. unadjusted disposable household income divided by the equivalence scale rate of group s. This says that a single-person household with income X enjoys the same standard of living as an s-person household with income XM_s. Buhmann et al. (1988) apply the similar approach $Y = \frac{X}{s^\theta}$ as equivalent income, resulting in the interpretation of θ, $0 \leq \theta \leq 1$, as an equivalence elasticity, with increasing economies of scale of household size in decreasing θ.

[6]As a follow-up of the primordial OECD scale, a new version of this scale, the modified OECD scale, attaches 1.0 to the leader, 0.5 to further adults above the age of 15 and 0.3 to children below the age of 15.

polarisation, whereupon chapter 5 subjects progression in taxation, and chapter 4 deals with the measurement of poverty and richness. In these four chapters, various indices of measurement are derived, compared to each other with respect to several aspects, sensitivities are discussed, applications presented, and advantages as well as disadvantages derived, so that results may be interpreted and the performance of indices evaluated. Chapter 6 finally concludes.

2   Measuring Inequality

After having identified an appropriate concept of income, one may apply analyses of various concepts of distributional effects. This section opens up with concepts of measuring traditional inequality at a distribution scale of incomes. Firstly, various indices of inequality are introduced, grouped by indices of dispersion, indices based on information theory, and normative indices. Then, these indices are compared to each other, with respect to fulfillment of fundamental axioms, with respect to sensitivity on the distribution scale, and with respect to sensitivity to equivalence scales.

2.1   Descriptive Measures / Measures of Dispersion

In the following, the main descriptive indices of inequality applied in econometric analyses are briefly introduced. These indices may be grouped as descriptive indices or measures of dispersion, since they only apply descriptive statistics in their calculus. The descriptive indices to be introduced are namely: the Gini coefficient, the relative mean deviation, the coefficient of variation, the logarithmic variance, the variance of the logarithms, the Mehran index, and the Piesch index. All these indices are based on a general concept of an index of inequality, presented in advance.

Let an income distribution for a homogeneous population consisting of n persons, with $n \geq 2$ be an equally distributed random variable $X = (x_1, x_2, ..., x_n)$, where $x_i \geq 0$ is the income of individual i, $i = 1, ...n$. Further on, X denotes a variable that may either be continuous or discrete and is defined on the interval $[a, b]$, with $a, b \, \epsilon \, \mathbb{R}$. The vector X is an element of D^n, the nonnegative orthant of the n-dimensional Euclidean space $\mathbb{R}^n$ without the origin, and the set of all income distributions is $D = \bigcup_{n \epsilon N} D^n$. Further, let $I : D \to \mathbb{R}$ be a continuous function, so that $I^m(X) \leq I^n(Y)$, with $m, n \, \epsilon \, \mathbb{N}$, $X \, \epsilon \, D^n$ and $Y \, \epsilon \, D^n$. Then, $I(\cdot)$ is called an *index of inequality*. Thus, each sequence $\{I^n : D^n \to \mathbb{R}^n\}_{n \epsilon \mathbb{N}}$ refers to a different population size n. In the case of a discrete variable, consider the ordered values $x_1 \leq x_2 \leq ... \leq x_i \leq ... \leq x_n$, or grouped values x_k, with $k = 1, ...m$, for $m \leq n$.[7] Following Piesch (1975)[8], a first index of dispersion may then be derived with the help of

---

[7]In case of X being continuous, $f(x)$ denotes the density function of X, $F(x)$ the corresponding strictly monotonously increasing distribution function, and $G[F(x)] = x$ the inverse distribution function defined on $[0, 1]$. The continuous case is of further subject in the appendix.

[8]Cf. Piesch (1975), pp. 37-39.

the concept of the Lorenz curve.

Firstly, cumulated absolute frequencies i are related to cumulated values s_i, i.e. $s_i = x_1 + ... + x_i$. Then, this summation function is standardized on $[0,1]$, and finally cumulated relative frequencies F_i are related to cumulated relative values L_i : $F_i = \frac{i}{n} \rightarrow L_i = \frac{s_i}{n\mu}$. The Lorenz curve then denotes all $n+1$ points generated by combining the two functions:

$$F_i = \frac{i}{n} = \sum_{j=1}^{i} \frac{1}{n} = \frac{\sum_{j=1}^{i} 1}{\sum_{j=1}^{n} 1} \text{ and } L_i = \sum_{j=1}^{i} \frac{x_i}{n\mu} = \sum_{j=1}^{i} l_i = \frac{\sum_{j=1}^{i} x_j}{\sum_{j=1}^{n} x_j}$$

with $l_i = \frac{x_i}{n\mu}, \forall \ i = 1, ... n.$[9] Since $F(x)$ and $L(x)$ are both distribution functions, the Lorenz curve is defined for the coordinate plane of the unit square, and always intersects the origin $(0,0)$ and the upper right corner of the unit square $(1,1)$. Moreover, since $L(x) \leq F(x)$, the Lorenz curve always runs beneath or at the straight diagonal of the unit square, where $L(x) = F(x)$. The Lorenz curve may be displayed as an increasing convex frequency polygon of n pieces running from $(0,0)$ to $(1,1)$, indicating how many percent of the sum of all values belong to the F-% smallest values of X.[10] Considering the area located between the Lorenz curve and the diagonal, the so called area of concentration A, and fractionalizing this area into segments of trapezoid-shape yields:

$$A = \frac{1}{2} \sum_{i=1}^{n} l_i \left(\frac{i-1}{n} + \frac{i}{n} \right) - \frac{1}{2} = \frac{1}{2} - \frac{1}{2} \sum_{i=1}^{n} \frac{1}{n} (L_{i-1} + L_i) = \sum_{i=1}^{n} l_i \frac{2i - n - 1}{2n}$$

displayed in figure 2.1.

Moreover, relating the area of concentration to the area of the triangle beneath the diagonal of the unit square, yields a first measure of dispersion for X being discrete[11], the so called **Gini coefficient** of inequality in its general version by Gini (1914)[12]:

$$I_{Gini}^{G} = \frac{A}{\frac{1}{2}} = 2A = \sum_{i=1}^{n} \frac{x_i}{n\mu} \frac{2i - n - 1}{n} = \sum_{i=1}^{n} l_i \frac{2i - n - 1}{n} = \frac{2 \sum_{i=1}^{n} i x_i}{n \sum_{i=1}^{n} x_i} - \frac{n+1}{n} \quad (2.1)$$

In case of maximum inequality, I_{Gini}^{G} corresponds to $1 - \frac{1}{n}$, and in the case of all values being equal, I_{Gini}^{G} corresponds to zero. One may derive a standardized Gini coefficient as

$$I_{Gini}^{*} = \frac{A}{\frac{1}{2}(1 - \frac{1}{n})} = \frac{n}{n-1} 2A = \frac{n}{n-1} I_{Gini}^{G} = \sum_{i=1}^{n} l_i \frac{2i - n - 1}{n - 1}$$

In case of maximum inequality, I_{Gini}^{*} corresponds to one, and in the case of all values

---

[9]The derivation of the Lorenz curve for the case of X being a continuous variable may be found in the appendix.

[10]Cf. Piesch (1975), p. 42.

[11]For X being continuous, it follows for the Gini coefficient: $I_{Gini}^{G} = \frac{A}{\frac{1}{2}} = 2A = 2(\frac{1}{2} - \int_0^1 L(F)dF) = 1 - 2\int_0^1 L(F)dF$, which is derived in further detail in the appendix.

[12]Cf. Gini (1914).
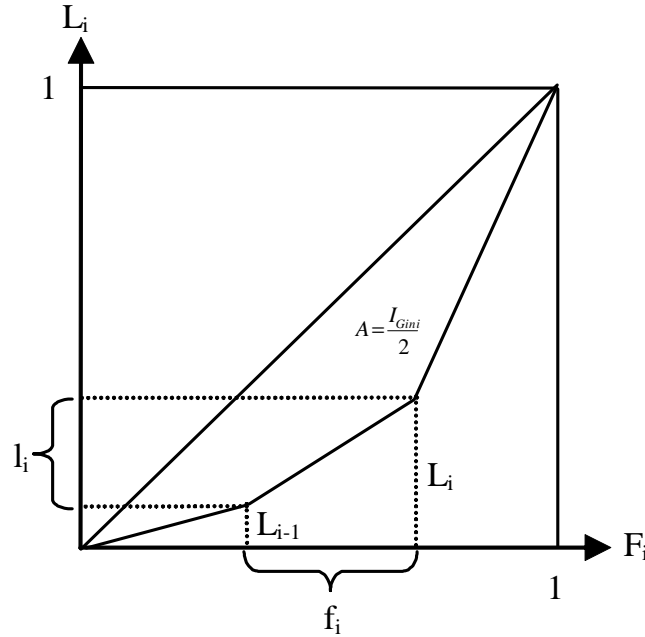
Source: Piesch (1975), p. 38.

Figure 2.1: The Gini Coefficient for a Discrete Variable

being equal, I^*_{Gini} corresponds to zero.[13]

Although the Gini coefficient became the probably most popular index of inequality in the latest decades, it is by far not the only index that has been applied in studies throughout literature, and it neither appears to be a perfectly appropriate index in all settings of analysis. For example, the Gini coefficient bears the drawback that it may indicate the same value of inequality for two distinct distributions in the case of intersecting Lorenz curves, since the Gini coefficient is a measure of overall dispersion, whereas it gives no information about dispersion in the upper or the lower level of the distribution. Therefore, other measures of dispersion should be presented in the following that might help solving this problem.

The most simple measure that considers the fact that values deviate from each other, is **the range**. It calculates the maximum spread of the distribution, i.e. $Range = x_{\max} - x_{\min}$. However, this measure takes only two values into consideration and neglects everything that happens between them.[14] In order to further elaborate the matter of deviation, one may apply **the relative mean deviation (RMD)**. The relative mean deviation is a measure that does not relate each value of X to each other, like the Gini coefficient does, it rather relates the deviation of each value x_i from the mean of the distribution, denoted by $\overline{x}$, to $\overline{x}$ itself. It follows

---

[13]Extentions of the Gini coefficient, alternative derivations, as well as a derivation for the case of a continuous variable can be found in the appendix.

[14]Cf. Cowell (1995), pp. 21-22.

$$RMD = \frac{\sum_{i=1}^{n} |x_i - \overline{x}|}{n\overline{x}} = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{x_i}{\overline{x}} - 1 \right| \tag{2.2}$$

in case of a discrete variable X.[15] RMD happens to correspond to the maximum deviation of the Lorenz curve from the diagonal line of absolute equality. i.e. $RMD = \max_{p\epsilon(0,1)}[p - L(p)]$.[16]

As usually when measuring dispersion of any frequency distribution, one may simply apply **the variance (VAR)** of the distribution:

$$VAR = \frac{1}{n} \sum_{i=1}^{n} (x_i - \overline{x})^2 \tag{2.3}$$

However, taking simply the variance as a measure of inequality yields the drawback that the degree of inequality is absolute, neglecting the mean around which the values spread. However, relating the variance to the mean of the distribution solves this problem and yields **the coefficient of variation (CV)** as another famous measure of inequality:

$$CV = \frac{\sqrt{V}}{\overline{x}} = \frac{\sqrt{\frac{1}{n} \sum_{i=1}^{n} (x_i - \overline{x})^2}}{\overline{x}} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (\frac{x_i}{\overline{x}} - 1)^2} \tag{2.4}$$

for X being discrete.[17] Another way of solving the problem of scale-variance is to take the logarithms of the values, i.e. $\log(x_i)$ and relating them to the logarithm of the mean income, i.e. $\log(\overline{x})$.[18] The resulting measure of inequality is called **the logarithmic variance (LVAR)**, denoted by[19]

$$LVAR = \frac{1}{n} \sum_{i=1}^{n} (\log \frac{x_i}{\overline{x}})^2 \tag{2.5}$$

The $LVAR$ is sometimes referred to as **the standard deviation of the logarithms**

---

[15] For a continuous variable, RMD denotes $\int_a^b \left| \frac{x_i}{\overline{x}} - 1 \right| f(x) dx$.

[16] RMD is also referred to as the Schutz coefficient or the Robin-Hood indicator, cf. Caminada and Goudswaard (2001), Annex A. When normalized on the interval $[0, 1]$, RMD is also called Kuznet's measure: $K = \frac{1}{2}RMD = \frac{1}{2n} \sum_{i=1}^{n} |x_i - \overline{x}|$, cf. Luethi (1981), pp. 30-33 and 91, or Pietra's measure: $P = \frac{1}{2n} \sum_{i=1}^{n} \left| \frac{x_i}{\overline{x}} - 1 \right|$, cf. Schmid and Trede (1999), p. 42.

[17] For for X being continuous, it follows: CV= $\sqrt{\int_a^b (\frac{x}{\overline{x}(F)} - 1)^2 dF(x)}$.

[18] One may also relate the $\log(x_i)$ to $\overline{\log(x_i)}$, instead of $\log(\overline{x})$, i.e. take the mean of the logs, instead of the log of the mean, which may lead to different results.

[19] Cf. Cowell (1995), pp. 20-21 and 24-25.

(SDL), which appears to be equivalent to the square root of the LVAR[20]:

$$SDL = \left[\frac{1}{n} \sum_{i=1}^{n} (\log \frac{x_i}{\overline{x}})^2 \right]^{\frac{1}{2}} = \sqrt{LVAR} \tag{2.6}$$

Taking the geometric mean $x^* = e^{\frac{1}{n} \sum_{i=1}^{n} \log x_i}$, instead of the arithmetic mean $\overline{x}$, the logarithm of the mean of incomes becomes $\log(x^*)$. Measuring inequality then yields the so called **variance of the logarithms (VARL):**

$$VARL = \frac{1}{n} \sum_{i=1}^{n} [\log(\frac{x_i}{x^*})]^2 \tag{2.7}$$

where it holds that $LVAR - VARL = [\log(\frac{\overline{x}}{x^*})]^2 > 0$.[21] Another index which belongs to a class of linear indices of inequality is defined by Mehran (1976) and is called **the Mehran index**:

$$I_{Mehran} = \frac{3}{n^3 \mu} \sum_{i=1}^{n} i(2n + 1 - i)(x_i - \mu) \tag{2.8}$$

Mehran (1976) also introduces a similar approach from the class of linear indices of inequality, referred to as **the Piesch index**[22]:

$$I_{Piesch} = \frac{3}{2n^3 \mu} \sum_{i=1}^{n} i(i - 1)(x_i - \mu) \tag{2.9}$$

2.2 Measures from Information Theory

Next to the group of indices of inequality that simply describe the distribution of a variable with respect to dispersion, another group of indices is derived with the help of a concept of probability of the occurrence of events that is based on information theory. Information theory focuses on messages about the occurrence of a specific event ω_i, out of the set Ω of possible events, with $P(\{\omega_i\}) = p_i$ denoting the probability that event ω_i will actually occur, $\sum_{i=1}^{n} p_i = 1$, $i = 1, ...n$. Before messages about the probability of occurrence come in, one may measure the expected information content of a message:

$$E(p_1, ...p_n) = \sum_{i=1}^{n} p_i e(p_i) = - \sum_{i=1}^{n} p_i \log p_i$$

with $e(p_i) = \log \frac{1}{p_i} = - \log p_i$ denoting the information content of a message. It is further defined: $0 \leq E \leq \log n$, with $E = 0$ if there is one i with $p_i = 1$, and all other $p_j = 0$ for

---

[20] The SDL measures inequality in case the variable may reasonably be assumed to have a lognormal distribution, since then the log-values of the variable have a normal distribution, the geometric mean is the median, the log geometric mean is the mean log, and the SDL measures dispersion. Only if the presumption of the lognormal distribution is valid, it holds that: $SDL = \sqrt{\ln(CV^2 + 1)}$.

[21] Cf. Newson (2004), pp. 388-391 and Cowell (1995), pp. 20-21 and 24-25.

[22] Cf. Piesch (1975).

$j \neq i$, i.e. minimum entropy, and $E = 1$ if $p_i = \frac{1}{n}, \forall\, i = 1, ...n$, i.e. maximum entropy.[23]

In Theil (1967) the entropy concept is applied to the measurement of inequality. He substitutes the probabilities p_i by income proportions $a_i = \frac{x_i}{n\bar{x}}$. In order to make the measure take its maximum value in case of maximum inequality, Theil (1967) subtracts entropy from its maximum value. Thus, inequality is measured by $\log n - E(a_1, ..., a_n)$. From this approach, he develops two measures: The first one corresponds to:

$$I_{Theil}^0 = \log n - \left(-\sum_{i=1}^{n} \frac{x_i}{\bar{x}n} \log \frac{x_i}{\bar{x}n}\right) = \log n + \sum_{i=1}^{n} \frac{x_i}{\bar{x}n} \log \frac{x_i}{\bar{x}n} = \frac{1}{n}\sum_{i=1}^{n} \frac{x_i}{\bar{x}} \log \frac{x_i}{\bar{x}} \qquad (2.10)$$

for X being discrete. I_{Theil}^0 is further on referred to as **the Theil index**. The second measure corresponds to:

$$I_{Theil}^1 = -\frac{1}{n}\sum_{i=1}^{n} \log \frac{x_i}{\bar{x}} = \frac{1}{n}\sum_{i=1}^{n} \log\left(\frac{\bar{x}}{x_i}\right) \qquad (2.11)$$

for X being discrete[24]. I_{Theil}^1 is also referred to as **the MLD**.[25]

In Shorrocks (1980) the entropy concept is also applied to measuring inequality. He introduces a class of inequality measures that deal with the extent to which inequality in the total population can be attributed to income differences between major population subgroups. He develops a generalization of Theil's approach of applying the entropy concept, i.e. **the indices of the generalized entropy (GE) family** of inequality:

$$I_{GE}^c(F) = \frac{1}{n}\frac{1}{c(c-1)}\sum_{i=1}^{n}\left[\left(\frac{x_i}{\bar{x}}\right)^c - 1\right], \quad -\infty < c < +\infty, \; c \neq 0, 1 \qquad (2.12)$$

for X being discrete[26]. The constant c is a sensitivity parameter, which may also be interpreted as a parameter of inequality aversion.[27] In case of $c = 0$, the indices of the GE family equal the **MLD**, i.e. I_{Theil}^1, in case of $c = 1$, they equal the Theil index, i.e. I_{Theil}^0, and in case of $c = 2$, they equal half the squared coefficient of variation, $I_{GE}^2 = \frac{CV^2}{2}$. In the case of $c = -1$, they are referred to as **the GE index**.

2.3 Normative Measures

Another group of indices of inequality is concerned with the concept of social welfare, which is closely related to the concept of entropy in information theory. The welfare analysis of distributional comparisons subjects individual preferences, uncertain prospects,

---

[23] Cf. Theil (1967), pp. 24-26 and Schmid and Trede (1999), p. 51.

[24] The Theil index corresponds to $I_{Theil}^0(F) = \int \frac{x}{\bar{x}(F)} \log \frac{x}{\bar{x}(F)} dF(x)$, for X being continuous, and the mean logarithmic deviation (MLD) to $I_{Theil}^1 = -\int \log\left(\frac{x}{\bar{x}(F)}\right) df(x)$, in this case.

[25] In case of $x_i = 0$, Theil defines $I(x_i = 0) = 0$. Cf. Theil (1967), pp. 93-95 and Schmid and Trede (1999), pp. 51-54.

[26] For X being continuous, $I_{GE}^c(F) = \frac{1}{n}\frac{1}{c(c-1)}\int_a^b\left[\left(\frac{x}{\bar{x}(F)}\right)^c - 1\right]dF(x)$.

[27] Cf. Shorrocks (1980), pp. 613-614 and 622.

coherent utility functions, the formulation of riskiness, and the concept of risk aversion. Thereby, social welfare functions build the link between welfare theory and inequality measurement, as they become a function of the equity of an income distribution. With the help of the concept of inequality aversion, it is assumed that social welfare increases the more equal incomes are distributed. The first approach of combining these theories goes back to Dalton (1920). He considers the average social welfare of the actual distribution of incomes

$$U_a = \frac{1}{n} \sum_{i=1}^n U(x_i) = \frac{1}{n(1-\varepsilon)} \sum_{i=1}^n x_i^{1-\varepsilon}$$

with $U' > 0$, $U'' \leq 0$, i.e. $U(x_i)$ is twice continuously differentiable, a function of x_i only, increasing in x_i and concave; moreover it is symmetric and additively separable in individual incomes. Thus, $U(x_i)$ may denote social utility or a welfare index of x_i, i.e. the social utility which the level and the rank of income x_i contributes to social welfare. It results the average social welfare U_a, as a *social welfare function*. U_a is then compared to the potential average social welfare that is achieved if all incomes are equal and equal the mean income of the actual distribution:

$$U_p = \frac{1}{n} \sum_{i=1}^n U(\overline{x}) = \frac{1}{1-\varepsilon} \overline{x}^{1-\varepsilon}$$

which yields **Dalton's Measure**:

$$I_D^\varepsilon = 1 - \frac{U_a}{U_p} = 1 - \frac{1}{n} \sum_{i=1}^n x_i^{1-\varepsilon} \overline{x}^{1-\varepsilon} \tag{2.13}$$

for $\varepsilon < 1$.[28]

However, the most famous approach stems from Atkinson (1970)[29], who further develops Dalton's approach. He applies the Lorenz curve, in order to compare two distributions, $f(x)$ and $f^*(x)$, where $F(x) \neq F^*(x)$ for some x and ranks them according to the following social welfare function:

$$W \equiv \int_0^{\overline{x}} U(x) f(x) dx$$

again with $U' > 0$, $U'' \leq 0$. He derives the result that one may judge on two distributions with the same mean value without further specifying $U(x)$, in the case that the Lorenz curves of the two distributions do not intersect, i.e. one can always find two functions that will rank the two distributions differently. Atkinson (1970) further concludes that two distributions can be ranked independently of the utility function if one distribution can be derived from the other by redistributing income from the richer to the poorer.[30] To be further able to make a complete ranking of distributions and quantify the degree of

---

[28] Cf. Dalton (1920) and Cowell (1995), pp. 46-47.

[29] Cf. Atkinson (1970), pp. 245-258.

[30] This result goes back to the fundamental Pigou-Dalton transfer principle, which is introduced later on. It demands that if the amount of money d from a person with income x_1 is transfered to a person with income x_2, in case $x_2 \leq x_1 - d$, the new distribution should always be preferred. For further derivation of the Pigou-Dalton transfer principle, see also the appendix.

inequality, Atkinson specifies $U(x)$ up to a monotonic linear transformation. He introduces a measure of inequality that is invariant with respect to linear transformations using the concept of the equally distributed equivalent level of income, x_{EDE}. That is the level of income per capita that is equally distributed among all individuals and at the same time yields the same level of social welfare as the original unequal distribution, i.e.

$$U(x_{EDE}) = \int_0^{\bar{x}} f(x)dx = \int_0^{\bar{x}} U(x)f(x)dx$$

A measure of inequality would then denote: $I_{EDE} = 1 - \frac{x_{EDE}}{\mu}$, with $0 \leq I_{EDE} \leq 1$, where $I_{EDE} = 0$ if incomes are distributed completely equally, i.e. $x_{EDE} = \mu$, and $I_{EDE} = 1$ if incomes are distributed completely unequally, i.e. $x_{EDE} = 0$. Thereby, inequality increases as I_{EDE} increases in x_{EDE}, i.e. $\frac{\partial I}{\partial x_{EDE}} = -\frac{1}{\mu}$. Thus, with increasing x_{EDE} everybody exhibits a higher equally distributed income.[31]

Atkinson (1970) further specifies the social utility function $U(x)$ with respect to the type of inequality aversion that characterizes the society. He proposes that people may feel more concerned about inequality with a rising average level of incomes, resulting in increasing relative inequality aversion in case of proportional additions to all incomes, and in an increasing I_{EDE}. If this is the case, the measure I_{EDE} may only be interpreted with reference to μ. Moreover he considers absolute equal additions θ_a to all incomes, which leads to absolute inequality-aversion, based on the development of the equally distributed equivalent income with respect to absolute changes in all income, $\frac{\partial x_{EDE}}{\partial \theta_a}$. Absolute inequality aversion increases if $\frac{\partial x_{EDE}}{\partial \theta_a} < 1$, it remains constant if $\frac{\partial x_{EDE}}{\partial \theta_a} = 1$, and it decreases if $\frac{\partial x_{EDE}}{\partial \theta_a} > 1$. Atkinson (1970) further derives the result that I_{EDE} may actually decrease, i.e. inequality decreases, with equal absolute additions to all incomes, even in case of increasing absolute inequality aversion. He finally adjusts the index I_{EDE}, in order to give the equally distributed equivalent income the property of invariance towards proportional translations to all incomes, deriving his famous **Atkinson index** of inequality:

$$I_A^\varepsilon = 1 - [\sum_{i=1}^n (\frac{x_i}{\mu})^{1-\epsilon} f(x_i)]^{\frac{1}{1-\varepsilon}}, \ \varepsilon \geq 0, \ \varepsilon \neq 1 \tag{2.14}$$

for X being discrete[32]. The parameter ε stands for the degree of inequality aversion, i.e. the relative sensitivity to transfers at different income levels. With ε increasing, more significance - concerning the degree of inequality - is attached to transfers at the lower end of the distribution scale and less significance to transfers at the top. As $\varepsilon \geq 0$, in the case of $\varepsilon = 0$ the utility function becomes linear and distributions are ranked solely according to total income, whereas in the case $\varepsilon \to \infty$, $U(x)$ is not strictly concave, taking account

---

[31] One may also interpret $\frac{x_{EDE}}{\mu}$ as percentage of the present national income that it costs the society to achieve - at equally distributed incomes - the same level of social welfare as is achieved at the present distribution, providing further implications for redistribution of incomes.

[32] For X being continuous, the Atkinson index denotes: $I_A^\varepsilon = 1 - \frac{1}{\mu(F)}[\int_a^b x^{1-\varepsilon} dF(x)]^{\frac{1}{1-\varepsilon}}$. Cf. Cowell (2000), p. 115.

only of transfers at the lowest income group. If $0 < \varepsilon < 1$, I_A^ε is ordinally equivalent to I_D^ε, specifically:

$$1 - I_D^\varepsilon = (1 - I_A^\varepsilon)^{1-\varepsilon}$$

for $\varepsilon \neq 1$[33], as it follows from equation 2.13 together with equation 2.14.[34] Moreover, for $\varepsilon = 1 - c$, it follows with equation 2.12 that I_A^ε is also ordinally equivalent to the indices of the GE family I_{GE}^c, specifically

$$I_{GE}^c = \frac{[1 - I_A^\varepsilon(F)]^c - 1}{c(c-1)}$$

The Atkinson index may therefore be interpreted as another special case of the GE family.[35]

2.4 Comparison of Indices

Before applying indices of inequality to the distributional analysis of a data set, one should compare them with respect to various characteristics, in order to point out specific advantages as well as weaknesses, differences as well as similarities, and thus assure an appropriate interpretation of their results.

2.4.1 Fulfillment of Axioms and Principles

In order to make their results comparable to each other, one may demand an index of inequality to fulfill several basic axioms and principles, some of which have already been mentioned[36]: an index of inequality fulfills the *axiom of monotonicity* if it indicates increasing inequality in case of a reduction in a low-level income and in case of an increase in a high-level income. The *axiom of normalization* demands the range of values of an index to be limited to $[0; 1]$. An index is *translation invariant* if inequality remains unchanged in turn of absolute as well as proportional translations to all incomes. The *axiom of symmetry* is fulfilled if inequality remains unchanged at any reordering of incomes, and the *population principle* demands that inequality remains unchanged if the population is replicated. An index is called *additively decomposable* if overall inequality may be decomposed into the sum of between-group inequality and within-group inequality, with the latter term being a weighted sum of the sub-group inequality values. The *Pigou-Dalton transfer principle* demands that a progressive transfer, i.e. a transfer from a richer to a poorer person that does not alter the relative ranks of the two, must always decrease

---

[33]In case of $\varepsilon = 1$, it is defined $I_A^\varepsilon = 1 - \frac{1}{\mu} \prod_{i=1}^n x_i^{\frac{1}{n}}$, which is also known to be the Champernowne measure of inequality. Cf. Chakravarty (1988), p. 152 and Luethi (1981), p. 50.

[34]Cf. Luethi (1981), pp. 51-52, Cowell (1995), pp. 46-47, and Dalton (1920).

[35]However, whether I_{GE}^c and I_A^ε are also cardinally equivalent, depends on the underlying social welfare functions. Cf. Cowell (2000), pp. 115 and 119.

[36]A methodological derivation of these axioms and principles as well as the derivation of the Lorenz dominance and the generalized Lorenz dominance criterion can be found in the appendix.

Index Nota-tion	Mono-tonicity	Normali-zation	Transl. Inv.	Sym-metry	Popu-lation	Decom-posabil.	Trans-fer P.	Dimin. Ret.
I_{Gini}^G	yes	on $[0;1]$	yes	yes	yes	no	yes	no
RMD	yes	on $[0;2]$	yes	yes	yes	no	no	no
VAR	yes	no	no	yes	yes	yes	yes	yes
CV	yes	no	yes	yes	yes	yes	yes	no
$LVAR$	yes	no	yes	yes	yes	no	no	no
$VARL$	yes	no	yes	yes	yes	no	no	no
MLD	yes	no	yes	yes	yes	yes	yes	yes
I_{Theil}^0	yes	no	yes	yes	yes	yes	yes	yes
I_{GE}^{-1}	yes	no	yes	yes	yes	yes	yes	yes
I_D^ε	yes	no	no	yes	yes	yes	yes	no
I_A^ε	yes	on $[0;1]$	yes	yes	yes	yes	yes	no

Table 2.1: Indices of Inequality - Fulfillment of Axioms and Principles

the degree of inequality, whereas a regressive transfer, i.e. a transfer from a poor to a richer person preserving relative ranks, must always increase the degree of inequality. An extension of this transfer principle - *the principle of diminishing returns* - assigns greater significance to a progressive transfer between two individuals with a given difference in incomes, if these incomes are low than if they are high, i.e. the magnitude of decrease in inequality is greater the lower are the incomes. The *principle of positional transfer sensitivity* demands that a transfer from any income to a lower one, with a fixed proportion of all incomes lying between these two, must have more significance at the lower end of the distribution scale than at the higher end.[37] The detailed performance of the various indices at fulfillment of these axioms may be found in the appendix. The results are only briefly summarized in table 2.1.

2.4.2 Sensitivity on the Distribution Scale

Due to an application of different underlying mathematical formulas, the various indices of inequality introduced so far vary greatly with respect to sensitivity to transfers along the distribution scale, even more than they vary with respect to fulfillment of axioms. While the results of some indices are relatively more sensitive to shifts among lower incomes, the results of other indices are relatively more sensitive to shifts among mid-level or high-level incomes. The indices of inequality are compared to each other with respect to this feature in the following.

The Gini coefficient is more sensitive in the lower levels of the income scale than in the higher levels, however it attaches the most weight to transfers among incomes in the middle of the scale. Thus it is most sensitive to transfers among mid-level incomes and generally in cases where values lie close to each other, and especially, in such cases it is highly sensitive compared to other indices. Similarly, **the relative mean deviation**, especially $\frac{1}{2}RMD$, is highly sensitive around the arithmetic mean income and relatively insensitive everywhere else. **The coefficient of variation** is more than average sensitive

---

[37]Cf. Dalton (1920), p. 351, Chakravarty and Muliere (2004a), pp. 8-12, and Kolm (1976), pp. 87-88.

among mid-level incomes and extremely sensitive to transfers in the highest level of the distribution scale, so that transfers of changes among the top 0.1% incomes often dominate the CV. The CV appears to be appropriate for the evaluation of transfers among mid-level incomes and especially the top of the income scale. **The Piesch index** is also relatively more sensitive to transfers among high incomes. However, **the logarithmic variance** and the **variance of the logarithms** are highly sensitive among low incomes, and they are more sensitive among mid-level incomes than the CV, I_{Theil}^0, and I_{Gini}^G.

Since it usually violates the Pigou-Dalton transfer principle in the upper level of the scale, the $LVAR$ appears to be only adequate for partial analyses in the middle and lower levels. The sensitivity of **the indices of the GE family** as well as **the Atkinson index** varies according to the value of their sensitivity parameters c and ε. For large absolute values of c, I_{GE}^c becomes more sensitive to variations in the tails of the distribution, specifically more sensitive in the upper scale for large positive values of c and more sensitive in the lower scale for large negative values of c. Thus, **the mean logarithmic deviation** is relatively more sensitive in the center, but also towards lower levels, while **the Theil index** is relatively medium-sensitive in mid- and low-levels and more than average sensitive in high-levels of the scale, but never as sensitive as the CV is in the highest levels.[38] Moreover, **the GE index** is relatively more sensitive in the lower levels, and $\frac{CV^2}{2}$ is relatively more sensitive in the upper levels. Generally, I_A^ε and I_D^ε are equally sensitive, especially equally sensitive as I_{Theil}^0 for $\varepsilon \to 0$, whereas for increasing ε they become more sensitive in the low-levels and less sensitive in the high-levels.[39] Table 2.2 summarizes the most important indices of inequality and their sensitivity on the distribution scale.

---

[38] This result is mainly based on Luethi (1981), while others certify the mean logarithmic deviation only more sensitivity in the lower-levels and the Theil index also more sensitivity in the lower-levels of the scale. It should though be noted that sensitivities of indices of inequality to transfers along the distribution scale may vary with respect to the underlying distribution of the income variable, as Luethi (1981) shows for the uniform distribution, the lognormal distribution, and the Pareto distribution.

[39] Cf. Atkinson (1970), pp. 255-257, Chakravarty (1988), p. 147, Luethi (1981), pp. 28-29 and 91-95, Buhmann et al. (1988), p. 125, and Champernowne (1974), p. 805.

Index			Sensitivity on the
Name	Notation	Formula	Distribution Scale
Gini coefficient	I_{Gini}^{G}	$\sum_{i=1}^{n} \frac{x_i}{n\mu} \frac{2i-n-1}{n}$	mid-level
Relative mean deviation	RMD	$\frac{1}{n} \sum_{i=1}^{n} \left\| \frac{x_i}{\overline{x}} - 1 \right\|$	mid-level
Variance	VAR	$\frac{1}{n} \sum_{i=1}^{n} (x_i - \overline{x})^2$	highest level
Coefficient of variation	CV	$\sqrt{\frac{1}{n} \sum_{i=1}^{n} (\frac{x_i}{\overline{x}} - 1)^2}$	highest level
Logarithmic variance	$LVAR$	$\frac{1}{n} \sum_{i=1}^{n} (\log \frac{x_i}{\overline{x}})^2$	low-level
Variance of the logarithms	$VARL$	$\frac{1}{n} \sum_{i=1}^{n} [\log(\frac{x_i}{x^*})]^2$	low-level
Mean logarithmic deviation	MLD	$\frac{1}{n} \sum_{i=1}^{n} \log(\frac{\overline{x}}{x_i})$	mid- and low-level
Theil index	I_{Theil}^{0}	$\frac{1}{n} \sum_{i=1}^{n} \frac{x_i}{\overline{x}} \log \frac{x_i}{\overline{x}}$	high-level
GE index	I_{GE}^{-1}	$\frac{1}{2n} \sum_{i=1}^{n} [\frac{\overline{x}}{x_i} - 1]$	low-level
Dalton index	I_{D}^{ε}	$1 - \frac{1}{n} \sum_{i=1}^{n} x_i^{1-\varepsilon} \overline{x}^{1-\varepsilon}$	high to low-level
Atkinson index	I_{A}^{ε}	$1 - [\sum_{i=1}^{n} (\frac{x_i}{\mu})^{1-\epsilon} f(x_i)]^{\frac{1}{1-\varepsilon}}$	high to low-level

Table 2.2: Indices of Inequality - Notation and Sensitivity on the Distribution Scale

2.4.3 Conclusion

All in all, when concluding on this whole chapter, the following most relevant differences between the indices of inequality derived so far shall briefly be summarized: it appears that the results of indices of inequality in empirical applications should be analysed with respect to relative changes in differences of the degree of inequality measured, rather than with respect to absolute values, since many indices are not normalized on $[0; 1]$. Moreover, most indices are sensitive at different ranges of the distribution scale, so that they implicitly measure different features of inequality at the same data set, and thus compute different absolute degrees of inequality, in case transfers are not distributed perfectly equal along the distribution scale. However, if one accounts for these differences in sensitivity with respect to the distribution scale, relative changes in differences in the absolute values of measures may be compared to each other, and cautious conclusions on the magnitude of changes in the degree of inequality may drawn. Moreover, adjustments for equivalence scales may be undertaken without any drawbacks, since indices of inequality then display which range on the distribution scale is mostly affected by such adjustments, if one again controls for the sensitivity of the indices with respect to transfers along the distribution scale.

3 Measuring Polarisation

The concept of polarisation has not been analyzed for long time yet in literature, since one has rather focused on inequality. Although inequality at an income distribution is generally reduced by income taxation systems, there may often be observed a development of two increasing peaks at the tails of distributions that move away from each other, creating a growing gap around the mean income. Such developments may lead to a social division into two groups, the very rich and the very poor, and are in recent literature referred to as polarisation of the income distribution. This section closely follows Schmidt (2004), since it is one of the latest extensive works on the measurement of polarisation and on the relation between polarisation, taxation and inequality.[40] Measures of polarisation may be grouped into two categories, i.e. measures based on axioms, and measures based on the concept of 'the declining middle class', introduced in the following.

3.1 Measures Based on Axioms

Following Schmidt (2004), the most simple and obvious indices of polarisation are also applied in various approaches of analysing inequality, in order to point out distances between certain ranges of the distribution scale. Such ratios apply the quantile function, denoted by

$$Q(F, q) = \min\{x \mid F(x) \geq q\} = x_q$$

at two distinct points of the distribution scale and compute the ratio of the values of $Q(F, q)$ at these points. Quantile ratios may be interpreted as the factor with that the incomes in the lower quantile in consideration need to be multiplied, in order to lift them up to the higher quantile, thus indicating a proportional gap between these quantiles. Specifically, mostly applied quantile-ratios are **the 0.75/0.25-quantile ratio**, also known as the quartile ratio:

$$PO_{QR}^{0.75/0.25} = \frac{Q(F, 0.75)}{Q(F, 0.25)} \tag{3.1}$$

and **the 0.9/0.1-quantile ratio**, also called the 0.9/0.1-decentile ratio:

$$PO_{QR}^{0.9/0.1} = \frac{Q(F, 0.9)}{Q(F, 0.1)} \tag{3.2}$$

with $Q(F, q)$ denoting the q-quantile. Two more of such ratios tell more about the absolute differences between all incomes in an upper quantile and all incomes in a lower quantile, in relation to the mean income of the overall distribution, i.e. **the 0.75/0.25-interquantile ratio:**

$$PO_{IQR}^{0.75/0.25} = \frac{Q(F, 0.75) - Q(F, 0.25)}{\mu_F} \tag{3.3}$$

---

[40]Cf. Schmidt (2004), pp. 5-41, 59-65, and 70-74.

and **the 0.9/0.1-interquantile ratio**:

$$PO_{IQR}^{0.9/0.1} = \frac{Q(F,0.9) - Q(F,0.1)}{\mu_F} \qquad (3.4)$$

One class of measures of polarisation is called the class of measures based on axioms, since indices in this class fulfill axioms that are similar to the axioms derived for indices of inequality.[41] All indices of this group are originally based on the fundamental Esteban-Ray index of income polarisation, derived by Esteban and Ray (1994). Following them, let $x_1, ..., x_n$ be values of a first variable, e.g. income X, that may be grouped into K disjoint groups according to a second variable, e.g. profession Y, with $\overline{x} = (\overline{x_1}, ..., \overline{x_K})'$ denoting the vector of mean incomes of the K groups, while $\overline{x_i} \neq \overline{x_j} \,\forall\, i,\, j$, i.e. mean incomes of two groups may never be equal. The vector of the K groups' fractions of the overall population is denoted by $w = (w_1, ... w_K)'$. Based on this categorization, Esteban and Ray (1994) characterize polarisation by the simultaneous occurrence of as well sufficiently large groups, denoting intra-group homogeneity as inter-group heterogeneity.

In a behavioural-economic model, *intra-group homogeneity* is applied by an *identification function*, $I : \mathbb{R}_+^K \to \mathbb{R}_+^K$, $w \to w^\alpha = (w_1^\alpha, ... w_K^\alpha)'$. Thereby, intra-group homogeneity increases in the degree of identification with people in the same group, which in turn increases in the number of people with the same income in this group and with decreasing differences between the incomes in the same group. Polarisation in turn increases in increasing intra-group homogeneity. Thereby, α is a parameter of polarisation sensitivity, with $1 \leq \alpha \leq 1.6$.[42] *Inter-group heterogeneity*, however, is applied by an *alienation function*, $V : \mathbb{R}_+^{K \times K} \to \mathbb{R}_+^{K \times K}$, $\overline{X} \to \left| \overline{X} - \overline{X}' \right|$, where alienation increases in increasing absolute differences between the mean incomes of the K groups. Polarisation in turn increases in increasing alienation. As a result, polarisation increases the more people with equal incomes belong to the same group and the greater are the differences between mean incomes of the groups. Based on this behavioural-economic model, Esteban and Ray (1994) derive the **Esteban-Ray index** of income polarisation as follows:

$$PO_{ER}^\alpha(\overline{X}, w) = \frac{1}{\overline{x}} (w^{1+\alpha})' \left| \overline{X} - \overline{X}' \right| w \qquad (3.5)$$

with , $\alpha \,\epsilon\, [1; 1.6]$ and $\overline{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$. It bears the advantages that it is based on a model approach with two specific partial functions, and that the differences compared to the measurement of inequality are revealed by a parameter of polarisation. Disadvantages of the index are the presumed a priori categorization into groups by a second variable and its representation by the groups' mean incomes, as well as the lack of representation of

---

[41] These axioms are the monotonicity axiom, the normalization axiom, the axiom of translation invariance, the symmetry axiom, the population principle, and the additive decomposability axiom. It should be noted that the Pigou-Dalton transfer principle is in its original version, as introduced at indices of inequality, not valid for the measurement of polarisation.

[42] The greater α, the greater is the difference between polarisation and inequality measured, whereat $\alpha = 0$ yields I_{Gini}^G.

deviation of incomes from the mean income *within* groups when regarding intra-group homogeneity, thereby overestimating polarisation. Moreover, the maximum value of PO_{ER}^{α} characterizes maximum inequality, instead of maximum polarisation. Nevertheless, the Esteban-Ray index is regarded one of the pioneer measures of polarisation.

In Esteban et al. (1999) PO_{ER}^{α} is expanded by a term that considers intra-group *in*homogeneity, a factor that is neglected by PO_{ER}^{α}. This additional term may be expressed by the difference between the Gini coefficient of the non-grouped income distribution and the Gini coefficient between the groups. The resulting **Esteban-Gradín-Ray index** of income polarisation corresponds to:

$$PO_{EGR}^{\alpha}(\overline{X}, w) = \frac{1}{\overline{x}} (w^{1+\alpha})' \left| \overline{X} - \overline{X}' \right| w - (I_{Gini}^{G} - I_{Gini}^{G,B}) = PO_{ER}^{\alpha} - I_{Gini}^{G,W} \qquad (3.6)$$

where I_{Gini}^{G} denotes the overall Gini coefficient, as defined in equation 2.1, $I_{Gini}^{G,B}$ denotes the Gini coefficient between the groups, $I_{Gini}^{G,W}$ denotes the Gini coefficient within the groups, and PO_{ER}^{α} is defined in equation 3.5. In order to minimize inequality within the groups, a statistic optimization tool allocates the incomes to the groups, thereby however maximizing inequality between the groups. In case of $\alpha = 0$, it follows that

$$PO_{EGR}^{0} = I_{Gini}^{G} - I_{Gini}^{G,W}$$

and in case of $\alpha = 1$, it holds that

$$PO_{EGR}^{1} = I_{Gini}^{G,B} - I_{Gini}^{G,W}$$

Generally, PO_{EGR}^{α} bears the following advantages: Firstly, the optimization tool makes it possible to measure polarisation independently of a second variable, and secondly, PO_{EGR}^{α} additionally considers intra-group *in*homogeneity. However, the generation of groups by an optimization tool contradicts the concept of identification in the behavioural-economic model of PO_{ER}^{α}. Moreover, this tool becomes complex for more than two groups, and it is left to open question according to which criterion the groups should be generated. Finally, similar to PO_{ER}^{α}, the maximum value of PO_{EGR}^{α} is characterized by maximum inequality, rather than maximum polarisation.

In Gradin (2000) PO_{EGR}^{α} is expanded by decomposing the overall population into partitions of subpopulations, thereby on the one hand yielding maximum polarisation in the overall population and on the other hand optimally describing a given degree of polarisation measured by PO_{EGR}^{α}. He builds subgroups according to a second variable, e.g. profession, and then measures group polarisation. His **Gradín index** denotes:

$$PO_{GRA}^{\alpha}(\overline{X}, w) = \frac{1}{\overline{x}} (w^{1+\alpha})' \left| \overline{X} - \overline{X}' \right| w - (I_{Gini}^{G} - I_{Gini}^{G,B,Sec} - 1) \qquad (3.7)$$

with $I_{Gini}^{G,B,Sek}$ denoting the Gini coefficient between the groups that have been generated

according to a second variable, whereat this decomposition follows the idea of finding the variable with maximum explained polarisation. This in turn is one of PO_{GRA}^α's advantages, next to the application of a correction term, $I_{Gini}^{G,B,Sec}$, that is economically reasonable, since it is based on a secondary variable. It remains again the same disadvantage, that PO_{GRA}^α with its maximum value characterizes maximum inequality, instead of maximum polarisation.

In D'Ambrosio (2001) on the one hand counterfactual kernel density estimates and on the other hand multiple secondary variables next to the primary variable income are applied, in order to differentiate within-group effects - the secondary variables held constant - from between-group effects. He then switches the secondary variable and thereby extracts its effects on overall income polarisation. He extends PO_{ER}^α by substituting the absolute distance between income means by the matrix of Kolmogorov's distances when incorporating inter-group heterogeneity. The **d'Ambrosio index** then follows as:

$$PO_{DAM}^\alpha(Kol, w) = (w^{1+\alpha})'wKol \tag{3.8}$$

with Kol denoting the matrix of Kolmogorov's distances[43]. PO_{DAM}^α bears the advantage that by decomposing the population non-parametrically with multiple secondary variables, specific factors and effects of polarisation may be identified and analysed separately. However, PO_{DAM}^α bears the drawbacks that the estimators for estimating the coefficient of overlapping, i.e. $1 - Kol = OLV_{kl}$, possess a substantial bias, and that estimating the asymptotic distribution as well as the standard error of PO_{DAM}^α is rather complex.

In Duclos et al. (2004) also the Esteban-Ray approach is expanded applying axioms for continuous distributions, while PO_{ER}^α is only valid for discrete distributions. They also apply non-parametrical kernel density estimates, in order to solve the problems of PO_{ER}^α and apply a behavioural-economic model. Identification of income x is denoted by the density function $f(x)$ and alienation is measured relatively to other incomes y, as $|x - y|$, rather than relatively to the mean income. The **Duclos-Esteban-Ray index** then results in:

$$PO_{DER}^\alpha(f) = \iint f(x)^{1+\alpha} f(y) |x - y| \, dxdy, \ \alpha \ \epsilon \ [0.25; 1] \tag{3.9}$$

With $\overline{I} = \int f(x)^{\alpha+1} dx$ denoting average identification and $\overline{V} = \int \int |x - y| \, dF(x)dF(y)$

---

[43]Kolmogorov's distances for two density functions f_k and f_l are defined as: $Kol_{kl} = \frac{1}{2}\int_{-\infty}^{+\infty} |f_k(x) - f_l(x)| \, dx$, yielding $Kol = \begin{bmatrix} Kol_{11} & ... & Kol_{K1} \\ ... & ... & ... \\ Kol_{1K} & ... & Kol_{KK} \end{bmatrix}$ as the matrix of Kolmogorov's distances. Then, $1 - Kol$ is defined as the coefficient of overlapping: $OVL_{kl} = 1 - \frac{1}{2}\int_{-\infty}^{+\infty} |f_k(x) - f_l(x)| \, dx = 1 - Kol_{kl}$.

denoting average alienation, PO_{DER}^{α} may be written as:

$$PO_{DER}^{\alpha} = \overline{IV}(1+\rho) = \overline{IV}(1 + \frac{Cov[I(X), V(X)]}{\overline{IV}}) = \overline{IV} + Cov[I(X), V(X)] \quad (3.10)$$

with $Cov[I(X), V(X)]$ denoting the covariance between identification and alienation. PO_{DER}^{α} bears the advantages that it is based on a behavioural-economic model, and that it is defined for continuous distributions. Moreover, the interaction between identification and alienation may be identified by the covariance, and instead of defining groups, identification is derived by non-parametric kernel density estimation. Disadvantages are that PO_{DER}^{α} is not invariant towards variations in the population, and its maximum value does characterize neither maximum polarisation nor maximum inequality. The kernel density estimation is rather complex, and PO_{DER}^{α} is rather ambiguous in α : while it decreases for $\alpha = 0.25$, it increases for $\alpha = 1$. However, it may be considered the most elaborate polarisation index based on axioms that is presented here.[44]

3.2 Measures and the Declining Middle Class

The concept of 'the declining middle class' more closely enlightens the differences between measuring inequality and measuring polarisation. While at the measurement of inequality the Pigou-Dalton transfer principle demands that a progressive transfer must always decrease inequality and a regressive transfer must always increase inequality, at the measurement of polarisation this principle is not valid. In order to derive this result, let income X be uniformly distributed on $[0;1]$, and make two progressive transfers that do not cross the median, with one above and one beneath the median. The graph of $f(x)$ clearly possesses two peaks then, i.e. $f(x)$ turned bimodal, thus polarisation clearly increases, while inequality decreases according to the Pigou-Dalton transfer principle. Thus, this principle is not valid for the measurement of polarisation. When focusing on 'the declining middle class', the following indices of polarisation highlight two matters characterizing polarisation, i.e. bimodality and spreadoutness. The first one characterizes a distribution with one mode above and one mode below the median income, while the latter one simply denotes deviation from the median income.

In Wolfson (1994) and Wolfson (1997) two polarisation curves are derived, in order to measure polarisation in the shade of the concept of 'the declining middle class'. Based on the empirical quantile function, they apply one major difference to the derivation of the Lorenz curve: the values are standardized by the median income m, instead of the mean income μ, yielding the *empirical quantile function of the median-standardized incomes*. This curve lies beneath the abscissa for values below the population fraction of 50% and above the abscissa for all values above 50%. Then mirroring the negative part of the empirical quantile function at the abscissa, yields Wolfson's *first polarisation curve*, displaying the deviation of the population fractions from the median income, which is

---
[44] Cf. Schmidt (2004), pp. 8-32 and 70-74.

I. Empirical Quantilefunction

Absolute Incomes

III. First Polarisation Curve

Absolute values of Median-standardized Incomes

II. Median-standardized empirical
Quantilefunction

Median-standardized Incomes

IV. Second Polarisation Curve

Kumul. absolute median-standardized Incomes
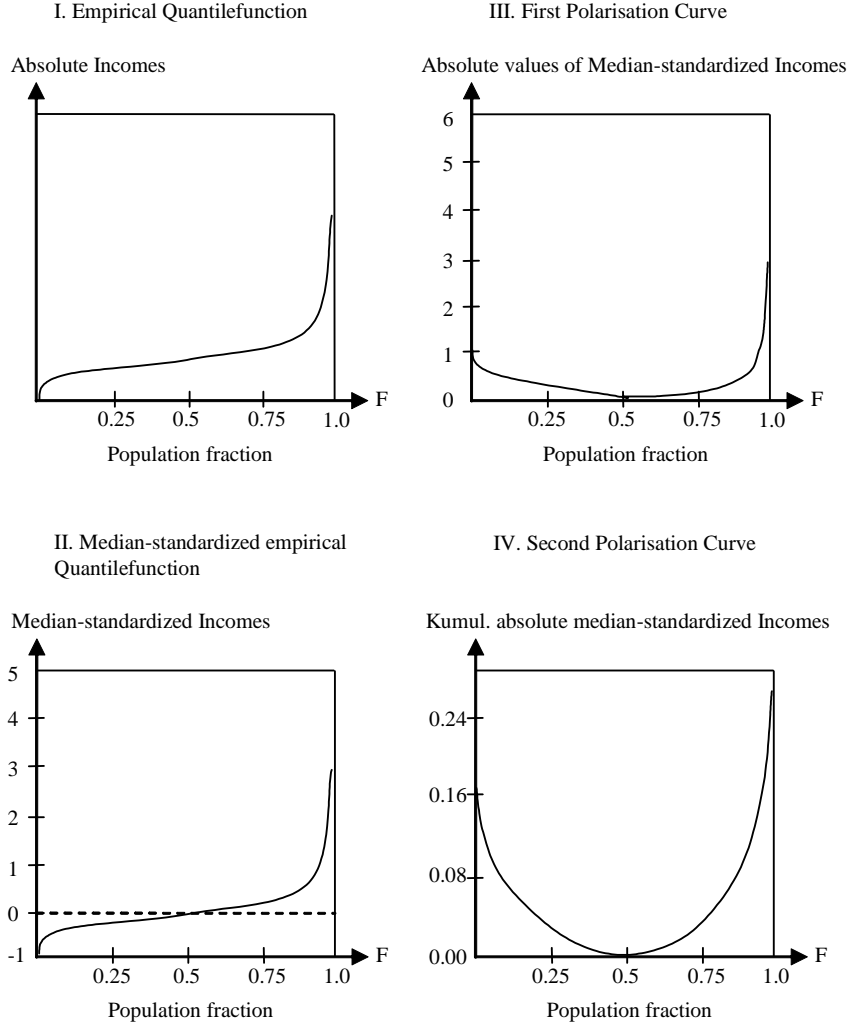
Source: Schmidt (2004), p. 61.

Figure 3.1: Derivation of Polarisation Curves

the central benchmark in the concept of 'the declining middle class'. Integrating the first polarisation curve in turn yields the *second polarisation curve*, which maps the cumulated deviations of the incomes from the median income. Figure 3.1 pictures the derivation of the polarisation curves.

Wolfson's index of polarisation, **the Wolfson index**, corresponds to four times the area beneath this curve, i.e.

$$PO_{WOL}(F) = 2\frac{\mu}{m}[1 - 2L_F(\frac{1}{2}) - I_{Gini}^G] \qquad (3.11)$$

with $L_F(\frac{1}{2})$ denoting the Lorenz curve at the 0.5-percentile, and I_{Gini}^G denoting the overall Gini coefficient.[45] PO_{WOL} bears the advantages that it links the measurement of

---

[45]This equation holds, since the ordinate of the second polarisation curve may be restandardized with $\frac{m}{\mu}$, yielding the ordinate-scale of the Lorenz curve, and then the abszissa may be shifted to fit the [1.0; 1.0]-plane diagonally, resulting in the Lorenz curve.

polarisation with the measurement of inequality in terms of the Lorenz curve, and that polarisation may be derived easily by the median tangent on the Lorenz curve, at the same time highlighting the differences to measuring inequality. Moreover, no groups need to be formed beforehand. PO_{WOL} bears the disadvantages that it is highly sensitive to the definition of the median income, especially in the case of few values, and it is not normalized on $[0; 1]$, rather it may take very high values in case of high inequality, when $\frac{\mu}{m}$ is very high.

In Wang and Tsui (2000) the approach of Wolfson (1994) is applied, characterizing polarisation by an increasing spread and by increasing bimodality. With the help of two axioms based on these properties, they derive a class of polarisation indices, **the Wang-Tsui index**:

$$PO_{WTS}^r(x) = \frac{1}{n} \sum_{i=1}^n \left| \frac{x_i - m}{m} \right|^r , \ r \ \epsilon \ [0; 1] \tag{3.12}$$

m denoting the median income and x_i denoting the increasingly ranked incomes. PO_{WTS}^r thereby measures the weighted relative deviation from the median income. Thus, PO_{WTS}^r is easily calculated and bears room for interpretation. Drawbacks include the fact that PO_{WTS}^r is highly sensitive to m, which thus needs to be clearly defined. PO_{WTS}^r is neither defined for maximum polarisation nor for $m = 0$. In some cases, PO_{WTS}^r yields contradicting results about an increase and a decrease in polarisation, and it takes values greater than one in case of high inequality.

In Rodriguez and Salas (2002) it is shown that the Wolfson index may be expressed by the Gini coefficient between the groups and the Gini coefficient within the groups:

$$PO_{WOL}(F) = 2\frac{\mu}{m}[1 - 2L_F(\frac{1}{2}) - I_{Gini}^G] = 2\frac{\mu}{m}[2(\frac{1}{2} - L_F(\frac{1}{2})) - I_{Gini}^G]$$

$$= 2\frac{\mu}{m}[2I_{Gini}^{G,B} - (I_{Gini}^{G,B} + I_{Gini}^{G,W})] = 2\frac{\mu}{m}(I_{Gini}^{G,B} - I_{Gini}^{G,W})$$

PO_{WOL} thereby measures the difference between inequality within the groups and inequality between the groups. Further applying an extended Gini coefficient,

$$I_{Gini}^{G,ext} = 1 - v(v-1)\int_0^1 (1-q)^{v-2} L_F(q)dq$$

they derive a general class of bipolarity measures, **the Rodríguez-Salas index**:

$$PO_{ROS}^v = I_{Gini}^{G,B(v)} - I_{Gini}^{G,W(v)}, \ v \ \epsilon \ [2; 3] \tag{3.13}$$

corresponding to the second polarisation curve by Wolfson (1997), with $I_{Gini}^{G,B(v)}$ denoting the Gini coefficient for between-group and $I_{Gini}^{G,W(v)}$ the one for within-group inequality. For $v = 2$, PO_{ROS}^v equals I_{Gini}^G. An advantage of PO_{ROS}^v is its result about the counteracting effects of as well $I_{Gini}^{G,B(v)}$, i.e. increasing polarisation, as of $I_{Gini}^{G,W(v)}$, i.e. decreasing

polarisation. Nevertheless, PO_{ROS}^v possesses all the drawbacks of PO_{WOL}.

Name	Notation	Formula
0.75/0.25-quantile ratio	$PO_{QR}^{0.75/0.25}$	$\frac{Q(F,0.75)}{Q(F,0.25)}$
0.9/0.1-quantile ratio	$PO_{QR}^{0.9/0.1}$	$\frac{Q(F,0.9)}{Q(F,0.1)}$
0.75/0.25-interquant. ratio	$PO_{IQR}^{0.75/0.25}$	$\frac{Q(F,0.75)-Q(F,0.25)}{\mu_F}$
0.9/0.1-interquantile ratio	$PO_{IQR}^{0.9/0.1}$	$\frac{Q(F,0.9)-Q(F,0.1)}{\mu_F}$
Esteban-Ray index	PO_{ER}^{α}	$\frac{1}{\overline{x}}(w^{1+\alpha})'\left|\overline{X}-\overline{X}'\right|w$
Esteban-Gardín-Ray ind.	PO_{EGR}^{α}	$\frac{1}{\overline{x}}(w^{1+\alpha})'\left|\overline{X}-\overline{X}'\right|w-(I_{Gini}^G-I_{Gini}^{G,B})$
Gradín index	PO_{GRA}^{α}	$\frac{1}{\overline{x}}(w^{1+\alpha})'\left|\overline{X}-\overline{X}'\right|w-(I_{Gini}^G-I_{Gini}^{G,B,Sec}-1)$
d'Ambrosio index	PO_{DAM}^{α}	$(w^{1+\alpha})'wKol$
Duclos-Esteban-Ray ind.	PO_{DER}^{α}	$\iint f(x)^{1+\alpha}f(y)\left|x-y\right|dxdy$
Wolfson index	PO_{WOL}	$2\frac{\mu}{m}[1-2L_F(\frac{1}{2})-I_{Gini}^G]$
Wang–Tsui index	PO_{WTS}^{r}	$\frac{1}{n}\sum_{i=1}^n\left|\frac{x_i-m}{m}\right|^r$
Rodríguez-Salas index	PO_{ROS}^{v}	$I_{Gini}^{G,B(v)}-I_{Gini}^{G,W(v)}$
Modified Wolfson index	PO_{WOL}^{mod}	$4[\frac{i}{n}-L_F(\frac{i}{n})-\frac{1}{2}I_{Gini}^G]$

Table 3.1: Indices of Polarisation - Notation and Formulas

In Schmidt (2004) PO_{WOL} is modified, in order to solve for its drawbacks, i.e. non-definition in case of a median income of $m=0$, very high values in case that the median is much smaller than the mean, and contradictory values in certain cases of progressive and regressive transfers. His **modified Wolfson index** denotes:

$$PO_{WOL}^{\mathrm{mod}}(F) = 2[2(\frac{i}{n}-L_F(\frac{i}{n}))-I_{Gini}^G] = 4[\frac{i}{n}-L_F(\frac{i}{n})-\frac{1}{2}I_{Gini}^G] = 2(I_{Gini}^{G,B}-I_{Gini}^{G,W}) \quad (3.14)$$

PO_{WOL}^{mod} bears the advantages that it solves for all the drawbacks of PO_{WOL}. But, since it equals twice PO_{EGR}^1 for $\alpha=1$, as it follows with equation 3.6:

$$PO_{WOL}^{\mathrm{mod}} = 2(I_{Gini}^{G,B}-I_{Gini}^{G,W}) = PO_{EGR}^1$$

it bears all of PO_{EGR}^{α}'s disadvantages as well.[46] Table 3.1 summarizes the indices of polarisation presented in this chapter.

---

[46]Moreover, Schmidt (2004) develops an own index of polarisation, based on the concept by Esteban and Ray (1994), that builds a direct link between the measurement of polarisation and inequality by deriving an index that is analogous to I_{Gini}^G and a curve that is analogous to the Lorenz curve. In addition his index is said to solve many of the other indices' problems. It should however not be of further subject in this analysis. Cf. Schmidt (2004), pp. 33-41 and 59-65.

4 Measuring Progression in Taxation

Many taxation systems follow the principle of progression, in order to redistribute incomes compared to the primordial distributions resulting from the power of markets. The idea behind redistribution is to use progression in taxation as a political device, in order to reduce the degree of inequality in an income distribution.[47] In this chapter, the concept of progression in taxation is introduced. Firstly, basic definitions and concepts are explained, whereupon local progression is differentiated from effective progression. Secondly, various indices of as well disproportionality as also redistribution are presented.

4.1 Definitions and Concepts

The fact whether a tax schedule is regarded progressive or not, is determined by as well the tax base as also the tax rate. According to taxation theory, income tax progression is generally characterized by an increasing average tax rate in percentage of income as income increases, i.e. the higher the income, the greater the *share* of this income that is paid for taxes. Thereby, progressive income taxation is accompanied by two effects, referred to as the *redistributive effect* on the one hand and *disproportionality*, also interpreted as deviation from proportionality, on the other.

Following Lambert (2001) and Schmidt (2004), let x be the income of a taxpayer and the twice differentiable function $t(x)$ denote the income tax schedule or tax liability, with $0 \leq t(x) < x$ and $0 \leq t'(x) < 1$. For strict progression it then holds:

$$\frac{d[\frac{t(x)}{x}]}{dx} > 0, \forall \ x > 0$$

while for weak progression it holds:

$$\frac{d[\frac{t(x)}{x}]}{dx} \geq 0, \forall \ x > 0$$

Let from now on

$$a(x) = \frac{t(x)}{x} \tag{4.1}$$

denote the *average tax rate* and

$$m(x) = \frac{dt(x)}{dx} = t'(x) \tag{4.2}$$

denote the *marginal tax rate*. Then, it follows from equations 4.1 and 4.2 that

$$\frac{d[\frac{t(x)}{x}]}{dx} = a'(x) = \frac{xt'(x) - t(x)}{x^2} = \frac{m(x) - a(x)}{x}$$

---

[47]In the following, it is presumed that tax liabilities are solely income-determined, i.e. other social non-income factors such as marital status, age and home-ownership are being neglected for the sake of simplicity.

and it follows for strict progression:

$$\frac{d[\frac{t(x)}{x}]}{dx} > 0 \Rightarrow m(x) > a(x), \forall\; x$$

i.e. the marginal tax rate lies everywhere above the average tax rate, for then they are both increasing and result in a strictly progressive tax system.

Let moreover x_0 denote an absolute amount of income that is exempted from taxation for political reasons. Such a tax exemption then furthermore differentiates *direct progression*, where there is no exemption granted, from *indirect progression*, exhibiting a positive tax exemption. Direct progression is characterized by an increasing marginal tax rate, i.e.

$$\frac{d[\frac{dt(x)}{dx}]}{dx} = t''(x) > 0$$

which in this case secures the progressive effect, whereas indirect progression is characterized by a constant marginal tax rate, i.e. $t''(x) = 0$, so that progressive effects result from the tax exemption only.[48]

The concept of indirect progression is the basis for the so called flat tax rate. Following Schmidt (2004), a *flat tax rate* may be defined by 1.) the tax liability:

$$t(x) = \max\{m(x - b); 0\}$$

with $b > 0$ denoting the tax exemption, $0 \leq m \leq 1$ denoting the constant marginal tax rate, and $x \geq 0$, together with 2.) the average tax rate:

$$a(x) = \frac{t(x)}{x} = \frac{\max\{m(x - b); 0\}}{x} = \begin{cases} m - \frac{mb}{x} & \text{if } x \geq b \\ 0 & \text{if } x < b \end{cases} \tag{4.3}$$

However, both types of progression, direct and indirect, exhibit an increasing average tax rate, so that they both possess progressive effects on the income distribution.[49]

A concept that is closely linked to progressivity is the concept of *redistribution*. The overall effects of redistribution of a tax system may be decomposed into two subeffects, the *vertical equity (VE) effect* and the *reranking (RR) effect*. While the concept of *horizontal equity* demands an equal tax treatment of taxpayers in identical circumstances, e.g. identical incomes, the concept of VE calls for an appropriate unequal treatment of unequals, i.e. unequal abilities of earning income, thereby enhancing redistribution.[50] However, if

---

[48] Cf. Lambert (2001), pp. 174-175, 187-189, and 196.

[49] In the case of a constant marginal tax rate and a positive absolute tax exemption, the average tax rate still increases, since the tax exemption's share of the total income declines in increasing income. Cf. Schmidt (2004), pp. 102-104.

[50] The ability-to-pay principle follows the concept of vertical equity when demanding a tax system to equalize everybody's loss in utility of income. Assuming a common increasing, twice differentiable and concave utility-of-income function $U(x), \forall\; x > 0$, this concept of equal loss in utility for all, i.e. $U(x) - U[x - t(x)] = u_0$, u_0 denoting an equal absolute reduction in utility, calls for progressive income taxation, rather than a proportional one, cf. Lambert (2001), pp. 174-175 and 183.

there appears reranking of incomes through taxation the net effect of redistribution of a tax system is counteracted. Thus,

$$L_{X-T}(p) - L_X(p) = C_{X-T}(p) - L_X(p) - [C_{X-T}(p) - L_{X-T}(p)] = VE - RR$$

i.e. redistribution expressed by the difference between the pre-tax and the post-tax Lorenz curves may be decomposed into the subeffect of VE and RR. The concept of *horizontal inequity* is closely linked to the effect of RR: RR of incomes by taxation is a necessary and at the same time sufficient condition for horizontal inequity. An index of horizontal inequity based on this concept will be introduced later on among the indices of redistribution.

It follows for the construction of an inequality-reducing progressive income tax system that such a system may demand all taxpayers to pay the same share of their income as taxes and still reduce inequality, i.e. implement a flat tax rate with a constant marginal tax rate, as long as there is a tax exemption granted at an appropriate level, i.e. an increasing average tax rate is guaranteed, in order to account for progressive effects.

4.2 Local versus Effective Progression

Further following Lambert (2001) and Schmidt (2004), so called measures of *structural progression*, also called *local progression*, measure the degree of income tax progression along the income scale, whereas so called measures of *effective progression* rather measure the degree of overall progression in a tax schedule's effects, given in a scalar index number. As shown above, for strict progression it must hold that $m(x) > a(x)$. Thus, a first reasonable index of local progression corresponds to **the first derivative of the average tax rate**:

$$PG_{AV}(x) = \frac{d[\frac{t(x)}{x}]}{dx} = \frac{xt'(x) - t(x)}{x^2} = \frac{m(x) - a(x)}{x} \tag{4.4}$$

PG_{AV} serves as a basis for two more important indices of local progression that measure the excess of the marginal tax rate over the average tax rate at income level x. The first one measures *liability progression*, defined as **the elasticity of tax liability to pre-tax income** at any x, with $t(x) > 0$:

$$PG_{LP}(x) = \varepsilon_{t(x),x} = \frac{dt(x)}{dx}\frac{x}{t(x)} = \frac{m(x)}{a(x)} \tag{4.5}$$

For a strictly liability progressive income tax system, it holds that

$$m(x) > a(x) \Leftrightarrow \frac{m(x)}{a(x)} > 1$$

i.e. a one per cent increase in pre-tax income leads to an increase in tax liability of more than one per cent. The second index measures *residual progression*, defined at any x as

the elasticity of post-tax income to pre-tax income:

$$PG_{RP}(x) = \varepsilon_{x-t(x),x} = \frac{d[x - t(x)]}{dx} \frac{x}{x - t(x)} = \frac{1 - m(x)}{1 - a(x)} \tag{4.6}$$

It indicates by which percentage the post-tax income increases if the pre-tax income increases by one per cent. For a residual progressive tax system it holds that $0 < \frac{1-m(x)}{1-a(x)} < 1$, i.e. the post-tax income increases by less than one per cent if the pre-tax income increases by one per cent. Moreover the degree of residual progression clearly increases if PG_{RP} decreases. Therefore it makes sense to define

$$PG_{RP}^*(x) = \frac{1}{PG_{RP}(x)} = \frac{1 - a(x)}{1 - m(x)}$$

so that it holds that the degree of residual progression increases with increasing PG_{RP}^*. Another index of local progression equals **the second derivative of the average tax rate**, which also measures the degree of local progression:

$$PG_{AV^2}(x) = \frac{d^2\left[\frac{t(x)}{x}\right]}{dx^2} = \frac{t''(x)}{x} - 2\frac{m(x) - a(x)}{x^2} \tag{4.7}$$

with $PG_{AV^2} > 0$ indicating accelerated progression, $PG_{AV^2} = 0$ indicating constant progression, and $PG_{AV^2} < 0$ indicating decelerated progression. For a flat tax rate, it follows with equations 4.4, 4.6, 4.5, and 4.7, in combination with equation 4.3, that:

$$PG_{AV}(x) = \begin{cases} \frac{mb}{x^2} & \text{if } x \geq b \\ 0 & \text{if } x < b \end{cases}$$

$$PG_{LP}(x) = \begin{cases} \frac{m}{m - \frac{mb}{x}} & \text{if } x \geq b \\ 0 & \text{if } x < b \end{cases}$$

$$PG_{RP}(x) = \begin{cases} \frac{1-m}{1-m+\frac{mb}{x}} & \text{if } x \geq b \\ 1 & \text{if } x < b \end{cases}$$

$$\text{and } PG_{AV^2}(x) = \begin{cases} -2\frac{mb}{x^3} & \text{if } x \geq b \\ 0 & \text{if } x < b \end{cases}$$

Musgrave, Thin (1948) introduce an index of effective progression which is independent of the local tax base, but rather considers the overall distribution of pre-tax and post-tax income:

$$PG_{MUT}^{eff}(x) = \frac{1 - I_{X-T}^{Gini}}{1 - I_X^{Gini}} \tag{4.8}$$

where I_X^{Gini} denotes the Gini coefficient of the pre-tax income distribution, and I_{X-T}^{Gini} denotes the Gini coefficient of the post-tax income distribution. Thereby, I_X^{Gini} and I_{X-T}^{Gini} are derived by simply applying I_{Gini}^G as defined in equation 2.1 to pre-tax incomes, as well as to post-tax incomes, respectively.

4.3 Indices of Disproportionality

This section introduces indices that are built on the concept of progressivity, which focuses on diversion from proportionality, based on Kakwani (1977)'s definition of progressivity as disproportionality. Thereby, a taxation schedule exhibits disproportionate effects if tax liabilities are not levied proportionately to incomes.[51] Such effects from progressive taxation may be shown again by applying the concept of the Lorenz curve. Next to tax liability $t(x)$, let $F(x)$ denote the distribution function of pre-tax incomes and $f(x)$ be its density function. Then it follows that

$$T(x) = n \int t(x)f(x)dx \tag{4.9}$$

may denote total revenue from income taxation, and

$$g(x) = \frac{T}{X} = \int \frac{f(x)t(x)dx}{\mu}, \ g > 0 \tag{4.10}$$

may denote the overall average tax rate or total tax ratio, with n denoting the number of all taxpayers and $X = n\mu$ the total pre-tax income. Then the Lorenz curve for pre-tax incomes follows as:

$$L_X(p) = \int_0^b \frac{xf(x)dx}{\mu} \tag{4.11}$$

and one may consider two other Lorenz curves, i.e. one for post-tax incomes:

$$L_{X-T}(p) = \int_0^b \frac{[x - t(x)]f(x)dx}{\mu(1 - g)} \tag{4.12}$$

and one for tax liabilities[52]:

$$L_T(p) = \int_0^b \frac{t(x)f(x)dx}{\mu g}, \ 0 \le p \le 1 \tag{4.13}$$

The difference $[L_X(p) - L_T(p)]$ may be interpreted as the fraction of the total tax burden shifted from low incomes, i.e. the bottom $100p$ per cent, to high incomes, i.e. the top $100(1 - p)$ per cent, by progression in the tax schedule. An index of disproportionality that is based on this difference is proposed by **Kakwani (1977)**:

$$PG_{KAK}(p) = 2\int_0^1 [L_X(p) - L_T(p)]dp \tag{4.14}$$

---

[51] Cf. Kakwani (1977).

[52] Precisely, $L_{X-T}(p)$ and $L_T(p)$ are concentration curves cumulating shares by rank. If assumed that no reranking occurs by taxation, they may be regarded as Lorenz curves, as Lambert (1994), p. 23 concludes.

Applying an extended Gini coefficient of the pre-tax income distribution

$$I_X^{Gini,ext}(v) = 1 - v(v-1)\int_0^1 (1-p)^{v-2} L_x(p)dp$$

and an extended concentration coefficient for tax liabilities

$$C_T^{ext}(v) = 1 - v(v-1)\int_0^1 (1-p)^{v-2} L_T(p)dp$$

an extension of PG_{KAK} can be derived as

$$PG_{KAK}^{ext}(v) = v(v-1)\int_0^1 (1-p)^{v-2}[L_X(p) - L_T(p)]dp = C_T(v) - I_X^{Gini,ext}(v) \qquad (4.15)$$

which focuses more on disproportionality towards the lower end of the income scale as v increases. Both, PG_{KAK} and PG_{KAK}^{ext}, increase if liability progression of an income tax increases at an unchanged pre-tax income distribution. Thus, they satisfy a consistency property, which states that at a given pre-tax income distribution, increasing local progression, in terms of liability progression, implies increasing effective progression, in terms of progressivity.

In Suits (1977) an index that is analogous to Kakwani's one is derived, in order to measure disproportionality, however, he builds it on relative concentration curves. Plotting cumulated fractions of tax liabilities on cumulated fractions of pre-tax incomes, yields the relative concentration curve of tax liabilities

$$C_T^{rel}(q) : q = L_X(p) \Rightarrow C_T^{rel}(q) = L_T(p) \qquad (4.16)$$

with $C_T^{rel}(q)$ being upward-sloping and convex for a progressive tax schedule. Then **Suits (1977)** measures aggregate disproportionality by:

$$PG_{SUI}(q) = 2\int_0^1 [q - C_T^{rel}(q)]dq = 2\int_0^1 [L_X(p) - L_T(p)]L_X'(p)dp \qquad (4.17)$$

Thus, PG_{SUI} can be obtained from PG_{KAK} by attaching the weight $L_X'(p)$ to the difference between the Lorenz curves, which then yields an index of effective progression. $PG_{SUI} \in [-1; +1]$, with $PG_{SUI} = -1$ in case of extreme regression, when the poorest pays all the taxes and $PG_{SUI} = 1$ in case of extreme progression, when the richest does so. However, $PG_{KAK} \in [-(1 + I_X^{Gini}); (1 - I_X^{Gini})]$, i.e. its boundaries depend on the degree of inequality in the income distribution, with $PG_{KAK} = -(1 + I_X^{Gini})$ in case of maximum regression and $PG_{KAK} = (1 - I_X^{Gini})$ in case of maximum progression.[53]

---

[53] Cf. Lambert (2001), pp. 201-204.

4.4 Indices of Redistribution

Analogously to considering the difference $[L_X(p) - L_T(p)]$ when measuring disproportionality, the difference $[L_{X-T}(p) - L_X(p)]$ may be interpreted as the fraction of the total post-tax income shifted from high incomes, i.e. the top $100(1 - p)$ per cent, to low incomes, i.e. the bottom $100p$ per cent, by progression in the tax schedule, indicating effects of overall redistribution of incomes. Moreover, analogously to quantifying disproportionality with PG_{KAK}, PG_{KAK}^{ext} and PG_{SUI}, **Reynolds and Smolensky (1977)** introduce an index measuring redistributive effects of progression based on the distance between $L_{X-T}(p)$ and $L_X(p)$:

$$PG_{RSM}(p) = 2\int_0^1 [L_{X-T}(p) - L_X(p)]dp \tag{4.18}$$

allowing a link to residual progression. With C_{X-T} denoting the concentration coefficient for post-tax incomes and I_X^{Gini} denoting the Gini coefficient of pre-tax incomes, it follows that[54]

$$PG_{RSM} = I_X^{Gini} - C_{X-T}$$

i.e. PG_{RSM} measures the reduction in the Gini coefficient resulting from the progressive tax schedule. Again analogously to PG_{KAK}^{ext}, there is:

$$PG_{RSM}^{ext}(v) = I_X^{Gini,ext}(v) - C_{X-T}^{ext}(v) \tag{4.19}$$

Similar to Suits (1977), **Pfaehler (1983)** extends PG_{RSM} to relative concentration curves, yielding:

$$PG_{PFA}(q) = 2\int_0^1 [q - C_{X-T}^{rel}(q)]dq = 2\int_0^1 [L_{X-T}(p) - L_X(p)]L'_X(p)dp \tag{4.20}$$

with C_{X-T}^{rel} denoting the relative concentration curve of post-tax incomes. Kiefer (1985) and Blackorby and Donaldson (1984) apply I_A^ε from equation 2.14, in order to measure redistributive effects of progression. **Kiefer (1985)** proposes an index similar to PG_{RSM}, i.e.

$$PG_{KIE}(\varepsilon) = I_{A,X}^\varepsilon - I_{A,X-T}^\varepsilon \tag{4.21}$$

whereas **Blackorby and Donaldson (1984)** measure the percentage increase in inequality as measured by I_A^ε:

$$PG_{BLD}(\varepsilon) = \frac{I_{A,X}^\varepsilon - I_{A,X-T}^\varepsilon}{1 - I_{A,X}^\varepsilon} \tag{4.22}$$

While two tax schedules, which make equal improvements in the Atkinson index through taxation, are judged equally progressive by PG_{KIE}, PG_{BLD} rates the schedule more

---

[54]Cf. Schmidt (2004), pp. 99 and 101-102, Musgrave and Thin (1948), as well as Lambert (2001), pp. 196-198.

progressive which possesses a more unequal pre-tax income distribution.[55]

An index introduced by **Blackburn (1989)** also quantifies redistributive effects of progression. Blackburn measures relative differences between inequality before and after taxation and concludes on redistributive effects. His **k-value** may be interpreted as the amount of money each taxpayer above the median income needs to pay to taxpayers below the median, in order to achieve exact equality between the pre-tax income distribution and the corresponding post-tax income distribution. When relating this k-value to the pre-tax mean income, it follows that[56]:

$$PG_{kBB}^{rel} = \frac{PG_{kBB}^{abs}}{\mu_X} = 2(I_{X-T}^{Gini} - I_X^{Gini}) \tag{4.23}$$

If one now wants to link the disproportionality effect of progression with its redistributive effect, one may apply the following relation between the Lorenz curves:

$$L_X = gL_T + (1-g)L_{X-T} \tag{4.24}$$

with $g = \frac{T}{X}$. Thus, the pre-tax Lorenz curve is a weighted average of the Lorenz curves of post-tax incomes and tax liabilities. It follows a significant result: If and only if tax liabilities are distributed more unequally than pre-tax incomes, i.e. higher incomes pay a greater share of their incomes for taxes than lower incomes, i.e. $L_T \leq L_X$, will income shares of given quantiles in the pre-tax distribution be more equally distributed after than before taxation, which is to say taxation redistributes incomes resulting in decreasing inequality, i.e. $L_{X-T} \geq L_X$. This result turns out to be a core characteristic of progressive income taxation, since it can be proven that

$$L_{X-T} \geq L_X \geq L_T \Leftrightarrow \frac{d[\frac{t(x)}{x}]}{dx} \geq 0, \forall \ F(X)$$

while in case of proportional taxation, it holds that

$$L_{X-T} = L_X = L_T$$

Moreover, it follows from transforming equation 4.24 into: $L_X - (1-g)L_X = gL_T + (1-g)L_{X-T} - (1-g)L_X$, and reordering it, that:

$$L_{X-T} - L_X = \frac{g}{1-g}(L_X - L_T) \tag{4.25}$$

with the left-hand side denoting redistribution and the right-hand side denoting disproportionality, weighted by the tax level $\frac{g}{1-g}$. Therefore, it follows with equations 4.14, 4.15,

[55] Cf. Lambert (2001), pp. 206-207 and 210.
[56] Cf. Merz et al. (2005), p. 8.

4.18, 4.19, 4.20, and 4.17 together with equation 4.25 that[57]:

$$PG_{RSM} = \frac{g}{1-g} PG_{KAK}$$

$$PG_{RSM}^{ext} = \frac{g}{1-g} PG_{KAK}^{ext}$$

as well as $PG_{PFA} = \frac{g}{1-g} PG_{SUI}$

Finally, an **index of horizontal inequity** is proposed by **Duclos and Araar (2006)**. From an index of horizontal inequity one would expect to quantify the degree of unequal treatment of equals, which appears to be a concept closely linked to the concept of redistribution, as introduced earlier. Hence the distance between the concentration curve of net incomes and the Lorenz curve of net incomes, which is an indicator of reranking, appears to be an appropriate index of horizontal inequity[58], i.e.

$$PG_{HI} = C_{X-T}(p) - L_{X-T}(p), \forall\ p\ \epsilon\ ]0,1[\tag{4.26}$$

All indices of progression in taxation systems introduced in this chapter and the corresponding concepts they focus measurement on, are summarized in table 4.1.

Name	Notation	Formula	Measure
First deriv. of avrg. tax rate	PG_{AV}	$\frac{m(x)-a(x)}{x}$	Loc. prg.
Elast. tax liabil./pre-tax inc.	PG_{LP}	$\frac{m(x)}{a(x)}$	Liab. prg.
Elast. post-tax/pre-tax inc.	PG_{RP}	$\frac{1-m(x)}{1-a(x)}$	Resi. prg.
Sec. deriv. of avrg. tax rate	PG_{AV^2}	$\frac{t''(x)}{x} - 2\frac{m(x)-a(x)}{x^2}$	Loc. prg.
Musgrave-Thin index	PG_{MUT}^{eff}	$\frac{1-I_{X-T}^{Gini}}{1-I_X^{Gini}}$	Eff. prg.
Kakwani index	PG_{KAK}	$2\int_0^1 [L_X(p) - L_T(p)]dp$	Disprop.
Suits index	PG_{SUI}	$2\int_0^1 [L_X(p) - L_T(p)]L'_X(p)dp$	Disprop.
Reynolds–Smolensky ind.	PG_{RSM}	$2\int_0^1 [L_{X-T}(p) - L_X(p)]dp$	Redistrib.
Pfaehler index	PG_{PFA}	$2\int [L_{X-T}(p) - L_X(p)]L'_X(p)dp$	Redistrib.
Kiefer index	PG_{KIE}	$I_{A,X}^{\varepsilon} - I_{A,X-T}^{\varepsilon}$	Redistrib.
Blackorby–Donaldson ind.	PG_{BLD}	$\frac{I_{A,X}^{\varepsilon} - I_{A,X-T}^{\varepsilon}}{1-I_{A,X}^{\varepsilon}}$	Redistrib.
Blackburn's k-value	PG_{kBB}^{rel}	$2(I_{X-T}^{Gini} - I_X^{Gini})$	Redistrib.
Index of horizontal inequity	PG_{HI}	$C_{X-T}(p) - L_{X-T}(p)$	Redistrib.

Table 4.1: Indices of Progression - Notation and Type of Progression Measured

---

[57] Cf. Lambert (2001), pp. 188-191 and 208-209.
[58] Cf. Duclos and Araar (2006).

5   Measuring Poverty and Richness

This chapter focuses on the low end of an income distribution. Firstly basic definitions and axioms for measures of poverty are briefly presented, then various indices measuring poverty are introduced and an evaluation of their fulfillment of axioms is presented. A rather new development, the measurement of richness is also shortly introduced.

5.1   Basic Definitions for Measures of Poverty

Following Lambert (2001), before one may measure any kind of poverty, one must make sure that it is precisely defined what one is about to measure. At poverty in particular, this means identifying who should be considered poor in the framework of the analysis and thereby making a clear cut at where poverty is determined to start. At the given data sets, it appears most appropriate to let an exogenously given poverty line determine this threshold. Although there are various socioeconomic variables feasible that may contribute relevant information to answering the question if a person may be considered relatively poor in given social surroundings, or not, again in the framework of this analysis the composition of the underlying data set limits the set of potential variables of poverty to individual incomes according to taxation statistics only, still leaving options for the differentiation between pre-government incomes and post-government incomes, as well as between unadjusted incomes and equivalence-scale adjusted incomes.[59] A poverty line then helps identifying the poor by representing the level of income necessary to maintain a subsistence level of standard of living. It may be defined either in absolute terms as a plain amount of pre-government or post-government income, adjusted or unadjusted, below which people are considered poor, or it may be defined relatively, e.g. to the mean or median income of the overall distribution. This leads to the classification of poverty indices as either absolute or relative poverty indices.

Let a poverty index be defined by a real valued function on $\mathbb{R}_+ \times \mathbb{Z}$, so that $PV^n(X, z)$ indicates the degree of poverty associated with any distribution of income $X \in \mathbb{R}_+^n$, with $z \in \mathbb{Z}$, $z > 0$ denoting the poverty line, and $n \in N$ denoting the overall number of incomes in the data set. A poverty index aggregates the characteristics of the poor into an indicator of poverty. Then PV^n is called a relative index of poverty if

$$PV^n(X, z) = PV^n(cX, cz) \ \forall \ n \in N, \ X \in \mathbb{R}_+^n, \ z \in \mathbb{Z}$$

where $c > 0$ is any scalar, whereas PV^n is called an absolute poverty index if

$$PV^n(X, z) = PV^n(X + c1^n, z + c) \ \forall \ n \in N, \ X \in \mathbb{R}_+^n, \ z \in \mathbb{Z}$$

---

[59]However, recent literature has shifted the emphasis of poverty analysis from the sole focus on individual income components to a multivariate focus on various attributes of well-being, like health, housing, environment, public goods, and literacy. Cf. Chakravarty and Muliere (2004b), pp. 28-30.

where c is a scalar such that $(X+c1^n) \epsilon \mathbb{R}^n_+$ and $(z+c) \epsilon \mathbb{Z}$. Moreover, one should in advance determine the intensity of poverty to be measured, e.g. how far people's incomes lie below a poverty line, and if this should be a relevant matter at all. When indices of poverty are introduced one will see that these matters mentioned here will be dealt with differently by the various indices. In the following, various indices of poverty are introduced, and their performance at the fulfillment of several axioms and principles, which are similar to the ones introduced at the measurement of inequality, is briefly evaluated.[60]

5.2   Indices of Poverty

In this section, various indices of poverty are introduced and evaluated with respect to fulfillment of the axioms presented. Chakravarty and Muliere (2004b) report an extensive list of indices which should be the base for the indices of poverty introduced in the following. Following them, let

$$Q(X) = \{i | x_i \leq z\}$$

denote the set of poor persons, x_i being person $i's$ income and q denoting the number of people having been identified as poor, according to their incomes at the income distribution of $X \epsilon \mathbb{R}^n_+$, i.e. the cardinality of the set $Q(X)$. Then, relating the number of poor people to the overall number of people in the population yields **the head-count ratio**, i.e.

$$PV_{HCR}(X, z) = \frac{q}{n} \tag{5.1}$$

It is as well an absolute index as a relative index. Furthermore, relating the average income shortfall of the poor to the poverty line, yields **the poverty-gap ratio**, i.e.

$$PV_{PGR}(X, z) = \frac{\sum_{i \epsilon Q(X)}(z - x_i)}{qz} \tag{5.2}$$

Combining equation 5.1 with equation 5.2 yields **the normalized poverty deficit**:

$$PV_{NPD}(X, z) = PV_{HCR}(X, z)PV_{PGR}(X, z) \tag{5.3}$$

Sen (1976) introduces an index that sums up the weighted income gaps among the poor, attaching higher weights to higher deprivation. His index became famous as **the Sen index**:

$$PV_{SEN}(X, z) = \frac{\sum_{i=1}^{q}(z - \widehat{x_i})(q + 1 - i)}{(q + 1)nz} \tag{5.4}$$

with $\widehat{x_i}$ denoting the illfare ordering of person i.[61] Blackorby and Donaldson (1980) introduce a generalization of the Sen index, i.e. **the Blackorby-Donaldson index** of

---

[60]More detailed definitions of these axioms as well as a more detailed analysis of the indices' performances at fulfilling them may be found in the appendix. Cf. Lambert (2001), pp. 133-136.

[61]For large q, the Sen index may be expressed by the head-count ratio, the poverty-gap ratio and the Gini coefficient among the poor: $PV_{SEN} = PV_{HCR}[PV_{PGR} + (1 - PV_{PGR})I^q_{Gini}]$.

poverty:

$$PV_{BLD}(X, z) = PV_{HCR}[1 - \frac{E^q(X^q)}{z}]$$ (5.5)

with $E^q(X^q)$ denoting the equally distributed equivalent income of the poor, evaluated according to a regular, homothetic social welfare function, and with PV_{HCR} from equation 5.1. PV_{BLD} measures the relative gap between the poverty line and the equally distributed equivalent income of the poor, times the number of poor people.[62] PV_{SEN} and PV_{BLD} are both sensitive to PV_{HCR}, to the degree of poverty among the poor, and to the degree of inequality among the poor.[63] Also generalizing the Sen index, Kakwani (1980) introduces **the Kakwani index** of poverty:

$$PV_{KAW}(X, z) = \frac{q}{nz \sum_{i=1}^{q} i^r} \sum_{i=1}^{q} (z - \widehat{x_i})(q + 1 - i)^r$$ (5.6)

for $r > 0$. For $r = 0$, it follows from equations 5.1 and 5.2 that

$$PV_{KAW} = PV_{HCR} PV_{PGR}$$

and for $r = 1$, from equations 5.6 and 5.4, it follows that the Kakwani index resembles the Sen index[64]:

$$PV_{KAW} = PV_{SEN}$$

In Hamada and Takayama (1977) an index that is based on a censored income distribution, X^*, replacing each non-poor income by the poverty line is introduced. **The Takayama index** then resembles the Gini coefficient of the poor:

$$PV_{TAK}(X, z) = \frac{1}{n^2 \mu(X^*)} \sum_{i=1}^{n} [2(n - i) + 1] \widehat{x_i^*}$$ (5.7)

In Chakravarty (1983) the proportionate gap between the poverty line and the equally distributed equivalent income $E^q(X^q)$ is applied, based on the censored income distribution. He derives **the Chakravarty index** of relative poverty[65]:

$$PV_{CHK}(X, z) = 1 - \frac{E^n(X^*)}{z}$$ (5.8)

Based on PV_{CHK}, Thon (1979) applies the rank of the poor persons in the total population as a weight of the income gap of the poor. **The Thon index** results in:

$$PV_{THO}(X, z) = \frac{2}{(n + 1)nz} \sum_{i=1}^{q} (z - \widehat{x_i})(n + 1 - i)$$ (5.9)

---

[62] An absolute version of PV_{BLD} corresponds to $PV_{BLD} = q[z - E^q(X^q)]$.

[63] Cf. Chakravarty and Muliere (2004b), pp. 18-28 and Sen (1976), pp. 223-225.

[64] Cf. Chakravarty and Muliere (2004b), pp. 20-21 and Kakwani (1980), p. 443.

[65] In the absolute version, PV_{CHK} denotes $PV_{CHK} = (z - E^n(X^*))$.

Clark et al. (1981) also build their **Clark, Hemming, Ulph (CHU) index** on PV_{CHK}. They apply the symmetric mean of order k for $E^n(X^*)$ in equation 5.8 yielding:

$$PV_{CHU}(X, z) = 1 - \frac{[\frac{1}{n}\sum_{i=1}^{n}(x_i^*)^k]^{\frac{1}{k}}}{z} \tag{5.10}$$

for $k < 1$, $k \neq 0$.[66] As k decreases greater weight is put to transfers at the lower end of the distribution.

In Foster and Shorrocks (1991) a group of subgroup decomposable indices is suggested. They define a continuous, decreasing and strictly convex function $f : \mathbb{R}^1_+ \rightarrow \mathbb{R}^1$, with $f(t) = 0$, $\forall\, t \geq 1$. **The Foster-Shorrocks indices** result in:

$$PV_{FSH}(X, z) = \frac{1}{n} \sum_{i\epsilon Q(X)} f(\frac{x_i}{z}) \tag{5.11}$$

For

$$f(t) = -\log t, t > 0$$

$PV_{FSH}^{n,1}$ becomes **the Watts index**, suggested by Watts (1968):

$$PV_{WAT}(X, z) = \frac{1}{n} \sum_{i\epsilon Q(X)} \log(\frac{z}{x_i}) = PV_{HCR}[I_{Theil}^q(X^p) - \log(1 - PV_{PGR})] \tag{5.12}$$

where I_{Theil}^q denotes the Theil index of inequality among the distribution of the poor incomes.[67] Foster et al. (1984) apply

$$f(t) = (1 - t)^\alpha, \alpha > 1$$

so that PV_{FSH} in equation 5.11 becomes **the Foster, Greer, Thorbecke (FGT) index**:

$$PV_{FGT}(X, z) = \frac{1}{n} \sum_{i\epsilon Q(X)} (\frac{z - x_i}{z})^\alpha \tag{5.13}$$

The coefficient α may be interpreted as a parameter of poverty aversion, since greater values of α attach increasingly greater weight to large poverty gaps.[68]

In Vaughan (1987) **the Vaughan index**, which measures the loss of welfare due to the presence of poverty, is introduced: It results in:

$$PV_{VAU}(X, z) = 1 - \frac{W^n(X)}{W^n(\widetilde{X})} \tag{5.14}$$

---

[66] In the case of $k = 0$, the CHU index denotes: $PV_{CHU}(X, z) = 1 - \frac{\prod_{i=1}^{n}(x_i^*)^{\frac{1}{n}}}{z}$.
[67] Cf. Chakravarty and Muliere (2004b), pp. 25-26.
[68] Cf. Chakravarty and Muliere (2004b), p. 27, Foster et al. (1984), pp. 762-764, and Coulter et al. (1992), p. 1071.

as a relative poverty index[69], where $\widetilde{X}$ is derived from X by setting all poor incomes equal to the poverty line, and $W^n(\cdot)$ denotes the underlying social welfare function. Finally Hagenaars (1987) extends the Vaughan index to **the Hagenaars index**, replacing X by X^* and assuming that the social welfare function corresponds to the sum of identical individual utility functions[70]:

$$PV_{HAG}(X,z) = 1 - \frac{1}{n} \sum_{i \epsilon Q(X)} \frac{U(x_i)}{U(z)} \tag{5.15}$$

All in all, the most popular indices of poverty that appear to be the most elaborate ones, are the following: the Sen index and the Kakwani index, which is built on the Sen index, moreover the Chakravarty index, and the CHU index, which is related to the Chakravarty index, and finally the Watts index and the FGT index which are both derived from the Foster-Shorrocks indices.

Name	Notation	Formula
Head-count ratio	PV_{HCR}	$\frac{q}{n}$
Poverty-gap ratio	PV_{PGR}	$\frac{\sum_{i\epsilon Q(X)}(z-x_i)}{qz}$
Normalized poverty deficit	PV_{NPD}	$PV_{HCR}(X,z)PV_{PGR}(X,z)$
Sen index	PV_{SEN}	$\frac{\sum_{i=1}^{q}(z-\widehat{x_i})(q+1-i)}{(q+1)nz}$
Blackorby-Donaldson index	PV_{BLD}	$PV_{HCR}[1 - \frac{E^q(X^q)}{z}]$
Kakwani index	PV_{KAW}	$\frac{q}{nz\sum_{i=1}^{q}i^r} \sum_{i=1}^{q}(z-\widehat{x_i})(q+1-i)^r$
Takayama index	PV_{TAK}	$\frac{1}{n^2\mu(X^*)}\sum_{i=1}^{n}[2(n-i)+1]\widehat{x_i^*}$
Chakravarty index	PV_{CHK}	$1 - \frac{E^n(X^*)}{z}$
Thon index	PV_{THO}	$\frac{2}{(n+1)nz}\sum_{i=1}^{q}(z-\widehat{x_i})(n+1-i)$
CHU index	PV_{CHU}	$1 - \frac{[\frac{1}{n}\sum_{i=1}^{n}(x_i^*)^k]^{\frac{1}{k}}}{z}$
Foster–Shorrocks index	PV_{FSH}	$\frac{1}{n}\sum_{i\epsilon Q(X)}f(\frac{x_i}{z})$
Watts index	PV_{WAT}	$\frac{1}{n}\sum_{i\epsilon Q(X)}\log(\frac{z}{x_i})$
FGT index	PV_{FGT}	$\frac{1}{n}\sum_{i\epsilon Q(X)}(\frac{z-x_i}{z})^{\alpha}$
Vaughan index	PV_{VAU}	$1 - \frac{W^n(X)}{W^n(\widetilde{X})}$
Hagenaars index	PV_{HAG}	$1 - \frac{1}{n}\sum_{i\epsilon Q(X)}\frac{U(x_i)}{U(z)}$

Table 5.1: Indices of Poverty - Notation and Formulas

All indices of poverty presented in this section are summarized in table 5.1, and their

---

[69]The absolute version of the Vaughan index corresponds to: $PV_{VAU}(X,z) = W^n(X) - W^n(\widetilde{X})$.
[70]Cf. Chakravarty and Muliere (2004b), pp. 27-28.

performance at the fulfillment of axioms, which are derived in the appendix, is summarized in table 5.2.

Index	Axioms and Principles									
Nota tion	Fo cus	Wk. mon.	Strg. trans.	Sym met.	Incr. pov.l.	Cont inuity	Popu lation	Mon. sen.	Dim.tr. sensit.	Subgr. Decom.
PV_{HCR}	yes	no	no	n/a	n/a	yes	n/a	no	no	yes
PV_{PGR}	n/a	yes	no	n/a	n/a	n/a	n/a	n/a	no	n/a
PV_{SEN}	yes	yes	no	yes	yes	no	no	n/a	n/a	no
PV_{BLD}	yes	yes	no	yes	yes	no	no	n/a	n/a	no
PV_{KAW}	n/a	n/a	no	n/a	n/a	no	no	n/a	yes	no
PV_{CHK}	yes	yes	yes	yes	yes	yes	y/n*	n/a	y/n*	y/n*
PV_{CHU}	yes	yes	yes	yes	yes	yes	yes	yes	yes	no
PV_{FSH}	yes	yes	yes	yes	n/a	yes	yes	yes	yes	yes
PV_{FGT}	yes	yes	yes	yes	n/a	yes	yes	yes	yes	yes
PV_{HAG}	n/a	yes	yes	n/a	yes	n/a	n/a	n/a	n/a	n/a

* Whether PV_{CHK} does or does not fulfill these axioms depends on the form of the underlying welfare function.

Table 5.2: Indices of Poverty - Fulfillment of Axioms and Principles

5.3 Measurement of Richness

While all poverty indices of the previous section are well-known, little research has been done on the measurement of richness.[71]

In a recent paper, Peichl et al. (2006) define a new class of richness measures. Let ρ be the richness line, *e.g.* 200% of median income, and $r = \#\{i|x_i > \rho, i = 1, 2, \ldots, n\}$ the number of rich persons. In most studies on income richness, only the proportion of rich persons is used as a measure of richness:

$$R_{HC}(\mathbf{x}) = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}_{x_i > \rho} = \frac{r}{n} \ .$$

Its definition resembles that one of the head count ratio for poverty. This definition of richness is not a satisfying, because this index will not change if nobody changes his or her status (rich or non-rich). Therefore, Peichl et al. (2006) define measures of richness R which are analogous to the FGT indices by

$$R(\mathbf{x}) = \frac{1}{n} \sum_{i=1}^{n} v\left(f\left(\frac{x_i}{\rho} \right) \right) ,$$

As the incomes of the rich have only a lower bound ρ, these incomes are transformed relative to the richness line, $\frac{x_i}{\rho}$, to the unit interval by a strictly increasing transformation function f. Where $f : \mathbb{R}_+ \rightarrow [0,1]$ is strictly increasing, $v : [0,1] \rightarrow \mathbb{R}_+$ (in particular

---

[71]For an overview of the sparse literature see Medeiros (2006).

$[0, 1]$) is increasing and $v(f(\cdot))$ is at least concave, that is, has a concave restriction on $[a, \infty[$ for some $a \in \mathbb{R}_+$.

Peichl et al. (2006) define $f(y) := 1 - \frac{1}{y}$ and $v(y) := y^\alpha$, with $\alpha > 0$, to obtain a richness index

$$R_\alpha(\mathbf{x}) = \frac{1}{n} \sum_{i=1}^n \left(1 - \frac{1}{\left(\frac{x_i}{\rho} \right) \mathbf{1}_{x_i > \rho}} \right)^\alpha = \frac{1}{n} \sum_{i=1}^n \left(\left(\frac{x_i - \rho}{x_i} \right)_+ \right)^\alpha, \qquad (5.16)$$

which resembles the FGT index.

One may also define $f(y) = 1 - \frac{1}{y^e}$, $e > 0$, for $y > 1$ and $v(y) = y$ and obtain an index similar to that one of Chakravarty[72]:

$$R_e(\mathbf{x}) = \frac{1}{n} \sum_{i=1}^n \left(1 - \left(\frac{\rho}{x_i} \right)^e \right)_+ , \ e > 0.$$

6 Conclusion

This paper provided a survey of the distributional analysis of fiscal reforms. Thereby, distributional effects have be differentiated by four subconcepts: inequality, polarisation, progression in taxation, and poverty and richness.

In order to properly prepare an analysis of the distributional effects of governmental activity, one might firstly want to adjust the income concept applied in the underlying data to specific effects of taxation law and derive a more market oriented concept of pre-government income. Further accounting for a heterogeneous population of tax units as well as differences in needs of households of different size by adjusting post-government disposable incomes to equivalence scales one is well equipped for an analysis of a data set, with respect to concepts of distributional effects, like inequality, polarisation, progression, and poverty.

When measuring inequality one may apply various indices, either descriptive indices simply measuring dispersion, or indices based on an entropy concept from information theory, or indices with a normative background of social welfare indices. Whichever indices applied, the results of empirical applications should be analysed with respect to relative changes in differences of the indices' values, rather than with respect to their absolute values, since many indices are not normalized. Moreover, most indices are sensitive at different ranges of the distribution scale, so that they implicitly measure different features of inequality at the same data set. However, if one accounts for these differences in

---

[72]See Peichl et al. (2006).

sensitivity, relative changes in differences of absolute values of measures may be compared to each other, and cautious conclusions on the magnitude of changes in the degree of inequality may be drawn.

Moreover, indices may be compared to each other with respect to fulfillment of certain desirable axioms and principles, as well as adjustments for equivalence scales may be undertaken without any drawbacks, since indices of inequality then mirror which range on the distribution scale is mostly affected by such adjustments, if one again controls for the sensitivity of the indices with respect to transfers along the distribution scale. Indices of inequality are also applied to empirical analyses, mainly with respect to their sensitivity along the distribution scale, with the Gini coefficient, the Atkinson index, and the Theil index being the most popular ones at this matter. Moreover, the Theil index appears to be a famous index for decomposition of overall effects into partial effects. It also appears to be highly useful, since its general class of GE indices can be adjusted to sensitivity towards all ranges of the distribution scale by its sensitivity parameter c. Also the Atkinson index appears to be useful, because of its normative character.

A subconcept of inequality subjects the formation of income groups which are moving away from each other on the income scale, exhibiting the development of two growing peaks at the tails of income distributions, creating a growing gap around the mean income. Such polarisation may be measured by on the one hand indices which are based on axioms, originating from simple quantile-ratios and the Esteban-Ray index, which are focused on the appropriate formation of groups and measure intra-group homogeneity applying an identification function as well as inter-group heterogeneity applying an alienation function, developing to the by now most elaborate Duclos-Esteban-Ray index, accounting also for interaction between identification and alienation. On the other hand, measures based on the concept of 'the declining middle class' focus on the growing gap between the two peaks on the distribution scale, generally all based on the Wolfson index.

Various concepts may be identified being responsible for a reduction in the degree of inequality and polarisation at income distributions. They all exhibit progressive effects of taxation systems, which may be grouped by redistributive effects from pre-tax to post-tax incomes, and effects of disproportionality at the determination of tax liabilities. The latter one is characterized by progressivity, which may further be decomposed into effects of vertical equity and reranking effects as well as horizontal inequity. Direct progression may be further differentiated from indirect progression according to the presence or the absence of an absolute tax exemption. Local progression may be differentiated from effective progression, whereat average tax rates may be compared to marginal tax rates, and elasticities may be calculated. Indices may be established as well. They generally consider the relation between certain Lorenz curves and concentration curves: while indices of disproportionality compare pre-tax Lorenz curves to tax-liability Lorenz curves, indices of redistribution compare pre-tax Lorenz curves to post-tax Lorenz curves. Relating these Lorenz curves to each other allows a link between indices of these two concepts

of progression, and thereby determine a progressive tax system.

When focusing on the low end of an income distribution, which is characterized by inequality and probably results from ongoing polarisation, one may, especially in countries other than the highly developed ones, find, although incomes are taxed progressively, the socially undesirable characteristic of poverty. If one then wants to measure a degree of poverty suffered among the very low incomes, one should beforehand make sure that it is precisely defined what one is about to measure, which means identifying who should be considered poor. Thereby an exogenously given poverty line might help determining an appropriate threshold. When limiting the analysis to individual income values from tax statistics, a poverty line appears helpful at identifying the poor by representing the level of income necessary to maintain a subsistence level of standard of living. Indices that may then be applied to the data, range from absolute ones simply counting heads below the poverty line, via relative ones accounting for poverty gaps, accounting for the mean among the poor incomes, attaching weights to higher deprivation, and determining equally distributed equivalent incomes. Indices again vary greatly at performance with respect to fulfillment of desirable axioms and principles. Also adjustment for appropriate equivalence scales makes sense, especially at the matter of poverty, where the absolute levels of incomes are of major relevance. The most popular indices of poverty, which also appear to be the most elaborate ones, are the following: the Sen index and the Kakwani index, which is built on the Sen index, moreover the Chakravarty index, and the CHU index, which is related to the Chakravarty index, and finally the Watts index and the FGT index, which are both derived from the Foster-Shorrocks indices.

All in all, it shall be concluded that the results on distributional effects of fiscal reforms are not as straightforward as popular phrases like 'the rich become richer, while the poor become poorer' and 'declining social justice' state. In general there occurs a need for a differentiated communication of the results of such analyses in political affairs, in order to prevent from one-sided and biased public perceptions of necessary public reforms. Only if it may be accomplished that voters have the abilities and at the same time the will to apply differentiated judgements on public reforms, long-term political power of such desirable reforms may be maintained.

Appendix 1. Basic Concepts of Measuring Dispersion - The Lorenz Curve

Many measures of inequality are based on the concept of the Lorenz curve which in turn is based on the concept of the distribution function and the inverse of the distribution function. This section closely follows Piesch (1975). He derives a class of measures of dispersion with the help of the inverse distribution function. Let X be a continuous variable defined on the interval $[a, b]$, with $a, b \in \mathbb{R}$. Then $f(x)$ denotes the density function of X, $F(x)$ the corresponding strictly monotonously increasing distribution function, and the inverse of it ,

$$G\left[F(x)\right] = x \tag{.1}$$

is called the inverse distribution function defined on the interval $[0, 1]$.[73] The arithmetic mean of $x = G(F)$ denotes

$$\mu = \int_0^1 G(F)dF$$

which appears to be the area beneath the inverse distribution function, and its variance becomes

$$\sigma^2 = \int_0^1 G^2 dF - (\int_0^1 GdF)^2$$

Furthermore, let

$$G^*(F) = \frac{G(F)}{\mu} \tag{.2}$$

denote the standardized inverse distribution function, which then posses an arithmetic mean of

$$\mu^* = \int_0^1 G^*(F)dF = 1$$

and

$$CV = \frac{\sigma^2}{\mu^2} = \int_0^1 G^*(G*-1)dF$$

is the so called coefficient of variation. Plugging the inverse distribution function in the general function for the means

$$\phi(k) = (\int_0^1 G^k(F)dF)^{\frac{1}{k}}$$

yields alternative means, i.e. the harmonic mean

$$h = (\int_0^1 \frac{dF}{G(F)})^{-1}$$

and the geometric mean

$$g = \lim(\int_0^1 G^k(F)dF)^{\frac{1}{k}})$$

---

[73]Piesch (1975) also defines the inverse distribution function for the case of a discrete variable as a left-continuous staircase function.

Considering the standardized inverse distribution function as a density function and then regarding the distribution function of this density function on $[0, 1]$, yields the so called Lorenz curve[74], the fundamental concept of measuring inequality.

Deriving the Lorenz curve, Piesch (1975) defines the first moment distribution of the continuous variable X as

$$L(x) = \frac{1}{\mu} \int_a^x u f(u) du$$

and calls it the incomplete first moment. The related distribution function $F(x)$ he calls the incomplete moment of order zero. Then he plots $F(x)$ and $L(x)$ on a rectangle coordinate plane[75] and derives the parametric notation of the Lorenz curve:

$$F(x) = \frac{\int_a^x f(u) du}{\int_a^x u f(u) du} = \int_a^x f(u) du = \int_a^x dF(u)$$

as the distribution function of X and

$$L(x) = \frac{\int_a^x u f(u) du}{\int_a^b u f(u) du} = \frac{1}{\mu} \int_a^x u f(u) du = \frac{1}{\mu} \int_a^x u dF(u)$$

as the moment distribution. Since $F(x)$ and $L(x)$ are both distribution functions, the Lorenz curve is defined for the coordinate plane of the unit square and always intersects the origin $(0, 0)$ as well as the upper right corner of the unit square $(1, 1)$. Moreover, since $L(x) \leq F(x)$, the Lorenz curve always runs beneath or at the straight diagonal of the unit square, where $L(x) = F(x)$. Plugging in equation .2 for x in the parametric notation yields:

$$L[G(F)] = L(F) = \frac{1}{\mu} \int_0^F G(F) dF = \int_0^F G^*(u) du$$

This is to say, values for the Lorenz curve resemble the fractional area beneath the standardized inverse distribution function. Thus, the Lorenz curve appears to be the distribution function of the standardized inverse distribution function. Its derivation is pictured in figure .1.

The Lorenz curve is monotonously increasing from $L(0) = 0$ to $L(1) = 1$, convex, and its first differentials in $L(0)$ and $L(1)$ correspond to $\frac{a}{\mu}$, and to $\frac{b}{\mu}$, respectively. Thus, if $a = 0$ the Lorenz curve starts with a horizontal tangent, and if $b = \infty$ it ends with a vertical tangent.[76] In order to be able to connect the single points of the Lorenz curve to a convex frequency polygon, Piesch (1975) applies the inverse distribution function as a staircase function and then yields $\frac{1}{\mu} \int_0^1 G(u) du$ for the Lorenz curve. However, if the Lorenz curve is approximated linearly between the points for each group, the dispersion calculated via the Gini coefficient may be underestimated. The Lorenz curve appears to be the fundamental method for any descriptive measure of dispersion.

---

[74] The term Lorenz curve refers to Lorenz (1905).

[75] In the case of a discrete variable, this would mean plotting the cumulated relative frequencies and the cumulated relative sums of the values of the variable.

[76] Cf. Piesch (1975), pp. 15-26.
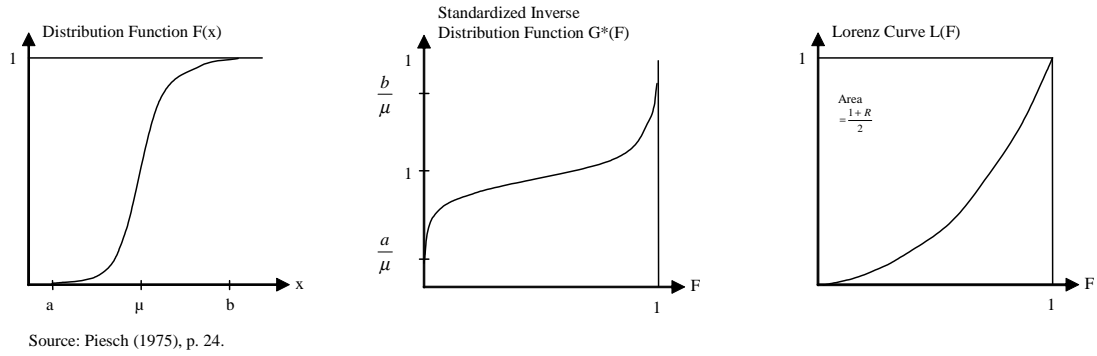
Source: Piesch (1975), p. 24.

Figure .1: Deriving the Lorenz Curve

Appendix 2. Descriptive Indices of Inequality - The Gini Coefficient

In case of X being a continuous variable, the Gini coefficient may be derived from the Lorenz curve. Following Piesch (1975), let

$$A = \int_0^1 d(F)dF = \frac{1}{2} - \int_0^1 L(F)dF, \ 0 \leq A \leq \frac{1}{2}$$

denote the so called area of concentration located between the diagonal and the Lorenz curve, with

$$d(F) = F - L(F)$$

being the difference function. Introduced by Gini (1914)[77], a measure of dispersion relates the area of concentration, A to the maximum area underneath the diagonal, as

$$I_{Gini}^G = \frac{A}{\frac{1}{2}} = 2A = 2(\frac{1}{2} - \int_0^1 L(F)dF) = 1 - 2\int_0^1 L(F)dF$$

and is called the Gini coefficient of inequality[78]. I_{Gini}^G can be interpreted as the mean of the difference function in relation to F on $[0,1]$, since

$$I_{Gini}^G = \frac{\int_0^1 d(F)dF}{\int_0^1 FdF}$$

Thereby, I_{Gini}^G is normalized on $[0,1]$, corresponding to zero in case of no dispersion, i.e. equality of all values, and one in case of maximum dispersion. Resulting, the area above

---

[77] Cf. Gini (1914).

[78] The 'G' indidates that this is the general Gini coefficient, in order to differentiate it from the more specific Gini coefficients for between-group and within-group inequality.
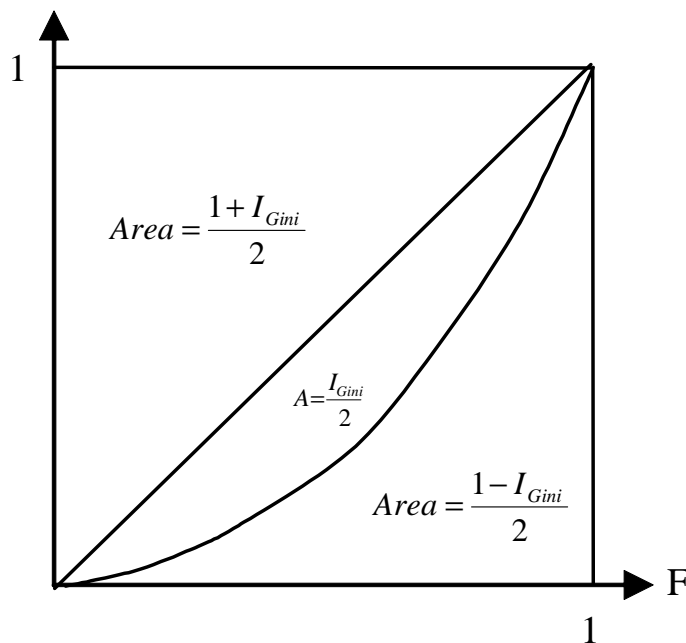
the Lorenz curve may be expressed by

$$\int_0^1 F(L)dL = \frac{1 - I_{Gini}^G}{2}$$

and the area beneath the Lorenz curve by

$$\int_0^1 L(F)dF = \frac{1 + I_{Gini}^G}{2}$$

where $L(F)$ denotes the inverse of the Lorenz curve. Figure .2 pictures the area of concentration and the Gini coefficient.

Source: Piesch (1975), p. 29.

Figure .2: Gini Coefficient for a Continuous Variable.

Moreover, the Gini coefficient can be calculated adapting equation .2, as

$$I_{Gini}^G = 2\int_0^1 FG^* dF - 1 = 2\int_0^1 FG^* dF - \int_0^1 G^* dF = 2\int_0^1 (2F - 1)G^* dF = 2cov(F, G^*)$$

Alternative notations of the Gini coefficient correspond to

$$I_{Gini}^G = 1 - 2\int_0^1 LdF = 2\int_0^1 FdL - 1 = 2\int_0^1 (1 - L)dF - 1 = \int_0^1 (1 - 2L)dF$$

and to

$$I_{Gini}^G = 2\int_0^1 FG^* dF - 1 = 1 - 2\int_0^1 (1 - F)G^* dF$$

The Gini coefficient may also be interpreted as a coefficient of variation, using Gini's

Mean Distance

$$\Delta = \iint_a^b |x - y| \, dF(x)dF(y)$$

Applying equation .1 it follows that

$$
\begin{aligned}
\Delta &= \iint_0^1 |G_x - G_y| \, dF_x dF_y = \int_0^1 \int_0^{F_x} |G_x - G_y| \, dF_x dF_y + \int_0^1 \int_{F_x}^1 |G_x - G_y| \, dF_x dF_y \\
&= \int_0^1 F_x G_x dF_x - \mu \int_0^1 L_x dF_x + \mu \int_0^1 (1 - L_y) dF_y - \int_0^1 G_x (1 - F_x) dF_x \\
&= 2 \int_0^1 FG dF - 2\mu \int_0^1 L dF = 2\mu \frac{1 + I_{Gini}^G}{2} - 2\mu \frac{1 - I_{Gini}^G}{2} = 2\mu I_{Gini}^G
\end{aligned}
$$

Thus,

$$I_{Gini}^G = \frac{\Delta}{2\mu} = \frac{1}{2} V_{(\Delta,\mu)}$$

i.e. the Gini coefficient corresponds to the coefficient of variation with Δ and μ, $V_{(\Delta,\mu)}$.[79]

Appendix 3. Definitions and Axioms for Measures of Inequality

Let an income distribution for a homogeneous population consisting of n persons, with $n \geq 2$ be an equally distributed random variable $X = (x_1, x_2, ..., x_n)$, where $x_i \geq 0$ is the income of individual i. The vector X is an element of D^n, the nonnegative orthant of the n-dimensional Euclidean space $\mathbb{R}^n$ without the origin, and the set of all income distributions is $D = \bigcup_{n \epsilon N} D^n$. Further, let the continuous function

$$I : D \to \mathbb{R}, \text{ so that } I^m(X) \leq I^n(Y), \text{ with } m, n \epsilon N, \ X \epsilon D^n \text{ and } Y \epsilon D^n$$

be an index of inequality. Thus, each sequence $\{I^n : D^n \to \mathbb{R}^n\}_{n \epsilon N}$ refers to a different population size n. Such an index of inequality may fulfill several axioms, introduced in the following.[80]

Monotonicity:

An inequality index fulfills the axiom of monotonicity if a reduction of a low-level income, ceteris paribus, unambiguously increases the degree of inequality, as a reduction of a high-level income unambiguously decreases inequality.[81]

Normalization:

An inequality index fulfills the axiom of normalization if its range of values is limited on the interval $[0; 1]$, i.e.

$$0 \leq I^n(X) \leq 1$$

---

[79] In the case of X being discrete, adapting similar transformations yields the same result for the two Gini coefficients and its relation to the coefficient of variation. With $\Delta = \frac{1}{n(n-1)} \sum_{i,j=1,i\neq j}^n |x_i - x_j|$ it follows that $I_{Gini}^* = \frac{\Delta}{2\mu}$ and $I_{Gini}^G = \frac{n-1}{n} \frac{\Delta}{2\mu}$. Cf. Piesch (1975), pp. 18-32 and 37-39.

[80] Cf. Chakravarty and Muliere (2004a), pp. 7-8.

[81] Cf. Foster et al. (1984), p. 762.

and if it holds that

$$I^n(X) = 0 \Leftrightarrow x_1 = x_2 = ... = x_n$$

i.e. the index takes the value of zero in case of equality of all incomes. Especially the upper limitation of 1 is violated by many famous indices of inequality.[82]

Translation Invariance:

An inequality index I corresponds to the concept of relative inequality if proportional translations to all incomes do not change inequality, i.e. $\forall\ n\ \epsilon\ N, \forall\ X\ \epsilon\ D^n$, $I^n(X) = I^n(cX)$, where $c > 0$ is a scalar, this is to say I is scale-invariant. In contrast, an index I is an absolute inequality index if it is invariant to equal absolute translations of incomes, i.e. $\forall\ n\ \epsilon\ N, \forall\ X\ \epsilon\ D^n$, $I^n(X) = I^n(X + c1^n)$, where c is a scalar so that $(X + c1^n)\ \epsilon\ D^n$.

Symmetry:

An inequality index I fulfills the axiom of symmetry if the degree of inequality remains unchanged under any reordering of incomes, i.e.

$$\forall\ n\ \epsilon\ N,\ \forall\ X\epsilon D^n,\ I^n(X) = I^n(Y)$$

where Y is any permutation of X, so that any two individuals may change their positions with no effect on inequality.[83]

Population Principle:

According to the population principle, inequality remains unchanged if a population is replicated m times:

$$\forall\ n\ \epsilon\ N,\ \forall\ X\ \epsilon\ D^n,\ I^n(X) = I^{mn}(Y)$$

where Y is the m-fold replication of X, i.e.

$$Y = (x^{(1)}, x^{(2)}, ..., x^{(m)})$$

with each $x^{(j)}$, $j = 1, ...m$ corresponding to X. The population principle is a property of all inequality indices that are defined on the continuum.[84]

Decomposability:

An inequality index is called decomposable if the inequality ranking between two distributions remains unchanged if both distributions are mixed with a third distribution, as long as all three distributions obtain the same mean value[85]:

$$\forall\ n\ \epsilon\ N,\ \forall\ X, Y, Z\ \epsilon\ D^n,\ \text{with}\ \mu_x = \mu_y = \mu_z$$

---

[82] Cf. Schmid and Trede (1999), pp. 36-37.
[83] Cf. Chakravarty and Muliere (2004a), pp. 8-9.
[84] Cf. Chakravarty and Muliere (2004a), p. 9 and Dalton (1920).
[85] Cf. Cowell (2000) and Cowell (1995), p.100.

it holds that

$$I^n(X) > I^n(Y) \Rightarrow I^n[(1-\delta)X + \delta Z] > I^n[(1-\delta)Y + \delta Z]$$

Furthermore, a decomposable index is called additively decomposable if overall inequality may be decomposed into the sum of between-group inequality and within-group inequality, with the latter term being a weighted sum of the sub-group inequality values.[86] Let there be $k = 1, ... K$ sub-groups, with inequality values I_k within these sub-groups and I_B capturing between-group inequality, i.e. differences in the general levels of income in the K groups, all within-group inequalities neglected. Then, overall inequality becomes

$$I = \sum_{k=1}^{K} \alpha_k I_k + I_B$$

with the weights

$$\alpha_k = \alpha(p_k, q_k)$$

depending on the proportion of overall population in group k, p_k, and group $k's$ share of total income, q_k, thus

$$\alpha : [0,1] \times [0,1] \rightarrow \mathbb{R}_+$$

The requirement of additive decomposability severely limits the set of adequate inequality measures in certain settings, where decomposition of overall effects on inequality shall be broken down to single partial effects.[87]

The Pigou-Dalton Transfer Principle:

Let $X \epsilon D^n$ be obtained from $Y \epsilon D^n$ by a so called progressive transfer if there exist two persons i and j such that

$$x_k = y_k \ \forall \ k \neq i, j; \ x_i - y_i = y_j - x_j > 0; \ y_i < x_i < y_j$$

and $y_i < x_j < y_j$. That is, X and Y are identical except for a positive transfer of income from person j to person i, with i having a lower income than j. Further, the transfer is such that it does not change the relative positions of the affected persons, i.e. it does not alter rank orders if the index fulfills the axiom of symmetry. Define a regressive transfer analogously so that equivalently Y is obtained from X by a regressive transfer. Then a famous property of inequality indices is the Pigou-Dalton transfer principle[88], demanding that a progressive transfer must always decrease the degree of inequality and a regressive transfer must always increase the degree of inequality:

$$\forall \ n \ \epsilon \ N, \ \forall \ Y \ \epsilon \ D^n, \ I^n(X) < I^n(Y)$$

---

[86] Cf. Shorrocks (1980), pp. 613-614.
[87] Cf. Lambert (2001), pp. 111-112.
[88] Cf. Dalton (1920), p. 351.

if X is obtained from Y by a progressive transfer.[89]

An extension of this principle - the principle of diminishing returns - assigns greater significance to a progressive transfer between two individuals with a given difference in incomes if these incomes are low than if they are high: $\forall\, n\, \epsilon\, N,\, \forall\, Y\, \epsilon\, D^n$, if X is obtained from Y by a progressive transfer from income $x_i + h$ to income x_i, $h > 0$, then

$$\frac{\partial\left[I^n(Y) - I^n(X)\right]}{\partial x_i} < 0$$

i.e. the magnitude of decrease in inequality is greater the lower is x_i.[90]

Appendix 4. Fulfillment of Axioms and Principles by the Indices of Inequality

The Gini coefficient is translation invariant, normalized on $[0, 1]$ and satisfies the axiom of symmetry as well as the population principle and the Pigou-Dalton transfer principle. However, it is not additively decomposable and violates the principle of diminishing returns, since it is linear in case of a given rank order of incomes.[91] Moreover, I_{Gini}^G is characterized by constant relative inequality-aversion. An extension of the Gini coefficient is introduced by Chakravarty (1988)[92]:

$$I_{Gini}^{G,ext} = 2\phi^{-1}[\textstyle\int_0^1 \phi(p - L(p))dp] \qquad (.3)$$

for X being continuous. $I_{Gini}^{G,ext}$ then fulfills the principle of diminishing returns.[93]

The relative mean deviation is translation invariant, fulfills the axiom of symmetry, the population principle, and is normalized on $[0, 2]$. One major disadvantage of RMD however is that it remains unchanged in case of transfers among incomes on only one side of the mean, since then the sum of absolute deviations from the mean does not change. Thus, it is not strictly concave and violates the Pigou-Dalton transfer principle,

---

[89] Cf. Chakravarty and Muliere (2004a), pp. 9-10.

[90] A combination of two transfers, a progressive one and a regressive one, such that the progressive one is taking place at a lower level of incomes than the regressive one, is called a favorable composite transfer if the variance of the original distribution does not change. Moreover, if X is obtained from Y by a favorable composite transfer and $I^n(X) < I^n(Y)$, this transfer has the property of transfer sensitivity. A positional version of the diminishing transfers principle is called the principle of positional transfer sensitivity demanding that a transfer from any income to a lower one with a fixed proportion of all incomes between these two must have more significance at the lower end of the distribution. See Kolm (1976), pp. 87-88 and Chakravarty and Muliere (2004a), pp. 10-12.

[91] When decomposing it into between-group inequality and within-group inequality a residual term occurs. Cf. Cowell (1995), p. 66, Maasoumi (1997), p. 211, Atkinson and Bourguignon (1987), p. 125, and Lambert (2001), p. 114. Lambert (2001) further uses the residual term for an evaluation of tax systems.

[92] He defines $\Psi(p) = \phi(p - L(p))$, $\phi : [0, 1] \to \mathbb{R}^1_+$ as a general divergence function, which is linear in $p - L(p)$. Assuming $\Psi(p)$ to be regular. A divergence function is called regular if it is continuous, strictly increasing, strictly convex and starts in the origin (0,0).

[93] Cf. Chakravarty (1988), pp. 148-149.

violates the principle of diminishing returns, and it is not decomposable.[94] However, it is characterized by constant relative inequality-aversion. **The variance** as an index of inequality satisfies the transfer principle and the principle of diminishing returns. It is decomposable, however not translation-invariant and not normalized on $[0, 1]$. Moreover, it is characterized by increasing inequality-aversion.[95]

The coefficient of variation is translation-invariant and decomposable, it fulfills the axiom of symmetry, the population principle and the Pigou-Dalton transfer principle. However, it violates the principle of diminishing returns, since it values transfers at high incomes way more than transfers at low incomes. Moreover, it is not normalized on $[0, 1]$, since it ranges on the interval $[0, \sqrt{n}]$, which indicates another disadvantage, i.e. that the CV ranges between very wide limits in case of many values. $\frac{CV^2}{2}$ is a special case of the GE family of inequality indices which possesses the property of additive decomposability, and it may therefore be used to analyze the effect of multiple components of income on inequality. The CV is characterized by constant relative inequality-aversion.[96]

The logarithmic variance is translation invariant, fulfills the axiom of symmetry and the population principle, however it violates the principle of diminishing returns, is not decomposable, and not normalized on $[0, 1]$. Moreover, it violates the Pigou-Dalton transfer principle among high incomes, i.e. for $x_i > e\overline{x}$, since then a progressive transfer increases rather than decreases the $LVAR$. Thus, the $LVAR$ seems to be inappropriate for an analysis at the upper level of the income scale.[97] However, it is characterized by constant relative inequality-aversion. **The variance of the logarithms** is translation invariant, however not decomposable and not normalized on $[0, 1]$. Moreover, it violates the Pigou-Dalton transfer principle and the principle of diminishing returns in the upper level of the income scale.[98]

The indices that belong to **the GE family** are translation invariant, satisfy the population principle, the Pigou-Dalton transfer principle and the principle of diminishing returns, however they are not normalized on $[0, 1]$. Moreover, these indices are additively decomposable with respect to subgroups, so that the GE family indices are also called the class of additively decomposable inequality measures. These indices allow for decomposability into the contribution due to differences between subgroups

$$B = \frac{1}{n} \sum_{g=1}^{G} n_g \log \frac{\mu}{\mu_g}$$

---

[94]Cf. Cowell (1995), pp. 22-23, 66, 139, and 142, Atkinson (1970), pp. 254-255, and Atkinson and Bourguignon (1987), p. 127.

[95]Cf. Cowell (1995), p. 66 and Atkinson (1970), p. 256.

[96]Cf. Cowell (1995), p. 66, Schmid and Trede (1999), p. 41, Atkinson and Bourguignon (1987), pp. 130-131, Kolm (1976), pp. 86-88, and Luethi (1981), pp. 34-35.

[97]Cf. Cowell (1995), p. 66, Atkinson and Bourguignon (1987), p. 127, and Schmid and Trede (1999), pp. 41-42.

[98]Cf. Cowell (1995), p. 66, Schmid and Trede (1999), p. 42, and Atkinson and Bourguignon (1987), p. 127.

and the contribution due to inequality within each subgroup $g = 1, ..., \dot{G}$:

$$C_g = \frac{1}{n} \sum_{i=1}^{n_g} \log \frac{\mu_g}{y_i^g}$$

The weights α_k for decomposition correspond to[99]

$$\alpha_k = \alpha(p_k, q_k) = p_k^{1-c} q_k^c$$

The Theil index fulfills all axioms and principles introduced above, except that it is not normalized on $[0, 1]$, since it has no upper limit. Especially, it satisfies the Pigou-Dalton transfer principle as well as the principle of diminishing returns. Most remarkably it is decomposable by components of income into $k = 1, ... K$ sub-group indices, according to decomposition of the indices of the GE family:

$$I_{Theil}^0 = \sum_{k=1}^K \left(\frac{n_k \overline{x}_k}{n \overline{x}}\right) I_{Theil}^{0,k} + \sum_{k=1}^K \frac{n_k}{n} \frac{\overline{x}_k}{\overline{x}} \log \frac{\overline{x}_k}{\overline{x}} \tag{.4}$$

with $I_{Theil}^{0,k}$ denoting the value of the Theil index of the k-th sub-group.[100]

    The Atkinson index is translation invariant and normalized on $[0, 1]$. Moreover, it is decomposable and satisfies the Pigou-Dalton transfer principle, whereas it violates the principle of diminishing returns. **The Dalton index**, however, is not translation invariant and not normalized on $[0, 1]$. Still, it also is additively decomposable, satisfies the Pigou-Dalton transfer principle and violates the principle of diminishing returns.[101]

Appendix 5. Ranking Distributions - The Lorenz Dominance Criterion

The matter of ranking distributions applies to normative indices. If one intends to not just specifically order some given distributions, but rather more generally to create a ranking over a set of distributions, one needs to extend the concepts of the standardized inverse distribution functions and the Lorenz curve, in order to derive dominance conditions. Dominance conditions may be desirable, since it can happen that multiple indices of inequality, when built upon several conflicting properties, come to a contradictory ranking

---

[99] Only in the cases of $c = 0$, i.e. the MLD, and of $c = 1$, i.e. the Theil index, do these weights sum up to 1. Moreover, in the first case the weights are independent of the income shares q_k so that within-group inequality simply corresponds to the sum of sub-group inequality weighted by population shares. Cf.Shorrocks (1980), pp. 619-621 and 625, Cowell (1995), p. 66, Shorrocks (1980), p. 625, and Lambert (2001), p. 112.

[100] It results an internal and an external inequality, similar to the decomposition of the variance. In case of a large internal inequality, groups are relatively inhomogeneous concerning incomes, and the inequality of the sub-group mean incomes only adds little to overall inequality. Cf. Bourguignon (1979), p. 915, Cowell (1995), p. 66, Theil (1967), pp. 93-95, and Schmid and Trede (1999), pp. 51-54.

[101] Cf. Cowell (1995), p. 66 and Bourguignon (1979), p. 918.

of incomes. To avoid such contradiction one may derive some dominance criteria.[102]

Consider a social welfare function

$$W \equiv \int_0^{\bar{x}} U(x) f(x) dx$$

with $U' > 0$, $U'' \leq 0$ generally, i.e. $U(x)$ is twice continuously differentiable, increasing and concave; moreover it is symmetric and additively separable in individual incomes. Further, allow for subclasses of W, where only $U' > 0$ holds and subclasses, where neither $U' > 0$, nor $U'' \leq 0$ hold. Consider the inverse distribution function

$$G\left[F(x)\right] = x$$

defined on the interval $[0, 1]$, where $F(x)$ denotes the corresponding strictly monotonously increasing distribution function, and $f(x)$ denotes the density function of X. The inverse distribution function resembles the quantile function, denoted by

$$Q(F, q) = \min\{x \mid F(x) \geq q\} = x_q$$

A distribution G then first-order distributionally dominates a distribution F, i.e. each quantile in G is no less than the corresponding quantile in F, if and only if

$$W(G) \geq W(F) \; \forall \; W$$

where at least $U' > 0$ holds. Let further

$$C(F, q) = \int_a^{Q(F,q)} x \, dF(x)$$

be the cumulative income function, where $C(F, 0) = 0$ and $C(F, 1) = \mu(F)$. Then, the graph of $C(F, q)$ against q is called the generalized Lorenz curve[103]. The generalized Lorenz curve can be derived from the conventional Lorenz curve by simply scaling it up by the mean:

$$GL(x) = C(F, q) = L(F, q)\mu(F)$$

so that now the vertical axis runs from 0 to μ, rather than from 0 to 1. A distribution G then second-order distributionally dominates a distribution F, i.e.

$$C(G, q) \geq C(F, q)$$

---

[102] The criteria introduced here all relate to the case of relative inequality indices. The case of absolute inequality indices is not elaborated here, since the number of such indices applied in literature appears to be very little.

[103] Cf. Shorrocks (1983), p. 6.

if and only if[104]

$$W(G) \geq W(F) \ \forall \ W$$

where both $U' > 0$ and $U'' \leq 0$ hold. This is equivalent to saying that G generalized Lorenz dominates F.[105] The result may be turned around, saying that the distribution G is at least as socially desirable as the distribution F, if and only if the generalized Lorenz curve of G lies at or above the generalized Lorenz curve of $F \ \forall \ 0 \leq q \leq 1$, i.e. especially that both curves do not intersect.[106]

Dividing the cumulative income function by the mean yields the conventional Lorenz curve, also known as the relative Lorenz curve:

$$\frac{C(F,q)}{\mu(F)} = \frac{1}{\mu(F)} \int_a^x u \, dF(u) = \frac{1}{\mu} \int_a^x u f(u) \, du = L(F,q) = L(x)$$

Then, a distribution G Lorenz dominates a distribution F, i.e.

$$L(G,q) \geq L(F,q)$$

if and only if

$$W(G) \geq W(F) \ \forall \ W$$

where again both $U' > 0$ and $U'' \leq 0$ hold. Only in case that $\mu_F = \mu_G$, i.e. the distributions have the same mean, do Lorenz dominance and generalized Lorenz dominance coincide. For conventional Lorenz dominance it can be stated: If distribution G Lorenz dominates distribution F, then G is regarded more equal then F by all inequality indices that fulfill the axiom of symmetry and the Pigou-Dalton transfer principle. Now, with conventional Lorenz dominance, conclusions about inequality comparisons are allowed.[107]

If one wants to take into account further non-income factors, such as family size, physical handicap, or location, one must assure for sequential dominance, as introduced by Atkinson and Bourguignon (1987), who made welfare comparisons in the presence of social heterogeneity. Their basic idea is to attribute different utility functions for monetary

---

[104]In primordial progression on this subject, also equal means for both distributions were assumed, cf. Atkinson (1970).

[105]Cf. Cowell (2000), Cowell (1995), pp. 100-105, and Chakravarty and Muliere (2004a), p. 28.

[106]This is also called the Shorrock's theorem, cf. Shorrocks (1983), p. 6. The case of intersecting generalized Lorenz curves, as further developed in Lambert (2001), p. 68. Thus, applying generalized Lorenz dominance one may rank distributions with differing means over a fixed population size, especially for the case that the conventional Lorenz curves cross. Moreover, the Sorrock's theorem allows comparisons, where there is Lorenz dominance but it is the smaller cake which is more equally distributed. It should however be noted that this ranking can only be one of social welfare, but not one of inequality. The criterion of generalized Lorenz dominance only allows judgements about the social welfare of two distributions. Whereas, to draw conclusions about the inequality of distributions one must stick to the following dominance criterion. Cf. Lambert (2001), p. 51.

[107]By introducing the principle of diminishing returns, Kolm (1976) shows which properties are necessary for inequality measures to possess, so that they assign greater significance to a progressive transfer between two individuals with a given difference in incomes if these incomes are low than if they are high. Cf. Cowell (2000), Cowell (1995), p. 105, Chakravarty and Muliere (2004a), pp. 25-26 and 28, and Kolm (1976), pp. 87-88.

income to different types of households. They then come to the result that a distribution G sequential dominates a distribution F, i.e. $W(G) \geq W(F)$, W being additive across all types of households and being subject to utility functions that are applied to different types, if and only if there is generalized Lorenz dominance of G over F in each of the sub-populations comprising the j most needy groups, $j = 1, ...n$.[108]

Appendix 6. Decomposability of the Coefficient of Variation

Let x_i be the overall income of person i, and let x_i be decomposable into K different income components,

$$x_{i1}, ..., x_{iK}, \text{ so that } \sum_{k=1}^{K} x_{ik} = x_i$$

Further, $Var(X)$ denotes the variance of the overall income, μ denotes the mean income, $Var(X_k)$ and μ_k denote the variance and the mean of income component k, respectively, $Cov(X_k, X_l)$ denotes the covariance, and ρ_{kl} denotes the correlation coefficient between income component k and l. CV^2 may then be decomposed into multiple income components:

$$
\begin{aligned}
CV^2 &= \frac{Var(X)}{\mu^2} = \sum_{k=1}^{K}(\frac{\mu_k}{\mu})^2 \frac{Var(X_k)}{\mu_k^2} + \sum_{k=1,l=1,k\neq l}^{K} \frac{\mu_k \mu_l}{\mu^2} \frac{Cov(X_k, X_l)}{\mu_k \mu_l} \\
&= \sum_{k=1}^{K}(\frac{\mu_k}{\mu})^2 CV_k^2 + \sum_{k=1,l=1,k\neq l}^{K} \frac{\mu_k \mu_l}{\mu^2} \rho_{kl} CV_k CV_l
\end{aligned}
$$

Thereby, the second term indicates the linear dependency of the multiple income components. Shorrocks (1982) introduces another decomposition of CV^2, focusing on the dependency between the multiple income components and overall income:

$$CV^2 = \frac{Var(X)}{\mu^2} = \sum_{k=1}^{K}(\frac{\mu_k}{\mu})^2 \frac{Cov(X_k, X)}{\mu_k^2}$$

Resulting from decomposition, one may conclude that an income component that makes up a large share of overall income causes a relatively large part of overall inequality.[109]

---

[108] Thereby, they start with the neediest group and one by one add the second neediest, checking at each stage for generalized Lorenz dominance till all groups are included, making the procecure one of sequential methods. This is also called the Atkinson-Bourguignon theorem. Cf. Atkinson and Bourguignon (1987), pp. 356-358.

[109] Cf. Schmid and Trede (1999), pp. 56-59 and Shorrocks (1982).

Appendix 7. Axioms and Principles for Indices of Poverty

Multiple axioms have been suggested in literature for relative as well as absolute poverty indices, the main ones correspond to the following - precise definitions follow this paragraph: *the focus axiom* suggested by Sen (1976) demands that a poverty index should be independent of the incomes of the non-poor. Sen also suggests that an index fulfills *the weak monotonicity axiom* if it indicates increasing poverty in the case that ceteris paribus there occurs a decrement in a poor person's income. Moreover, it fulfills *the strong transfer axiom* if ceteris paribus there happens a regressive transfer from a poor person to someone who is richer. *The axiom of symmetry* demands that a poverty index is invariant to any reordering of incomes. If a poverty index unambiguously measures an increasing degree of poverty, in case the poverty line is shifted upwards, the index fulfills *the increasing poverty line axiom* suggested by Clark et al. (1981), and if the index is a continuous function of incomes, the index fulfills *the continuity axiom*. If the degree of poverty measured remains unchanged if the population is replicated, the index fulfills *the population principle axiom* introduced by Chakravarty (1983) and Thon (1983).[110] Kakwani (1980) suggests that a poverty index should be relatively more sensitive to transfers among the very poor incomes than among any other incomes and introduces two axioms: an index fulfills *the monotonicity sensitivity axiom* if it indicates a greater increase in poverty due to a decrement in a poor person's income, the poorer the person is. Kakwani (1980) also suggests *the diminishing transfer sensitivity axiom* for an index to fulfill, if it indicates a greater increase in poverty due to a regressive transfer from a poor person with income x_i to another poor person with income $x_i + h$, $h > 0$, the lower is x_i, with none of the two poor crossing the poverty line due to this transfer.[111] Finally, Foster et al. (1984) suggest *the subgroup decomposability axiom* for an index of poverty if the indicated overall degree of poverty may be decomposed into various degrees of poverty, attributed to subgroups that may be formed by partitioning the population by some homogeneous characteristic.[112]

The definitions of these axioms and principles correspond to the following:

The focus axiom:

$$PV^n(X, z) = PV^n(Y, z) \ \forall \ n \ \epsilon \ N, \ X, Y \ \epsilon \ \mathbb{R}^n_+, \ z \ \epsilon \ \mathbb{Z}$$

if

$$Q(X) = Q(Y)$$

and if

$$x_i = y_i, \forall \ i \ \epsilon \ Q(X)$$

---

[110] Cf. Chakravarty and Muliere (2004b), pp. 9-13 and Sen (1976).
[111] Cf. Kakwani (1980), pp. 438-439.
[112] Cf. Chakravarty and Muliere (2004b), pp. 13-18 and Foster et al. (1984), pp. 763-764.

with

$$Q(X) = \{i | x_i \leq z\}$$

being the set of poor persons, and x_i being person $i's$ income.

The weak monotonicity axiom:

$$PV^n(X, z) < PV^n(Y, z)$$

in the case that Y is obtained from X by a decrement in a poor person's income.

The strong transfer axiom:

$$PV^n(X, z) < PV^n(Y, z)$$

with Y being obtained from X by a regressive transfer from a poor person to someone who is richer.

The axiom of symmetry:

$$PV^n(X, z) = PV^n(Y, z)$$

if Y is obtained from X by a permutation of incomes.

The increasing poverty line axiom: $PV^n(X, z)$ is increasing in z.

The continuity axiom: $PV^n(X, z)$ is a continuous function of X.

The population principle axiom:

$$PV^n(X, z) = PV^{mn}(Y, z)$$

where Y is the m-fold replication of X, i.e.

$$Y = (X^{(1)}, ..., X^{(n)})$$

with each $X^{(i)}$ being X.

The monotonicity sensitivity axiom:

$$PV^n(Y^1, z) - PV^n(X, z) > PV^n(Y^2, z) - PV^n(X, z), \ \forall \ Y^1, Y^2 \ \epsilon \ \mathbb{R}^n_+$$

obtained from X by the same amount of decrement to poor incomes x_i and x_j, with $x_i < x_j$.

The diminishing transfer sensitivity axiom:

$$PV^n(Y, z) - PV^n(X, z)$$

is greater the lower x_i, if Y is obtained from X by a regressive transfer from a poor person with income x_i to a poor person with income $x_i + h$, $h > 0$, none of the two poor crossing

the poverty line due to this transfer.[113]

The subgroup decomposability axiom:

$$PV^n(X, z) = \sum_{i=1}^{m} \frac{n_i}{n} PV^{n_i}(X^i, z),$$

for $X^i \in \mathbb{R}_+^n$, $i = 1, 2, ..., m$; $X = (X^1, X^2, ..., X^m)$, and $\sum_{i=1}^{m} n_i = n$

---

[113]Cf. Chakravarty and Muliere (2004b), pp. 9-18, Kakwani (1980), pp. 438-439, and Foster et al. (1984), pp. 763-764.

Bibliography

Atkinson, A.B. (1970). On the Measurement of Inequality, *Journal of Economic Theory* **2**: 244-263.

Atkinson, A.B. (1987). On the Measurement of Poverty, *Econometrica* **55:** 749-764.

Atkinson, A.B. and Bourguignon, F. (1987). Income Distribution and Differences in Needs, in G.R. Feiwel (ed.). Arrow and the Foundations of the Theory of Economic Policy, pp. 350-370, Macmillan, New York.

Blackburn, McKinley L. (1989). Interpreting the Magnitude of Changes in Measures of Income Inequality, *Journal of Econometrics* **42**: 21-25.

Blackorby, C. and Donaldson, D. (1980). A Theoretical Treatment of Measures of Absolute Inequality, *International Economic Review* **21**: 107-136.

Blackorby, C. and Donaldson, D. (1984). Ethical Social Index Numbers and the Measurement of Effective Tax/Benefit Progressivity, *Canadian Journal of Economics* **17**: 683-694.

Bohnet, A. (1999). Finanzwissenschaft: Grundlagen Staatlicher Verteilungspolitik, 2nd edition, Oldenbourg, Munich.

Bourguignon, F. (1979). Decomposable Income Inequality Measures, *Econometrica* **47**: 901-920.

Buhmann, B., Rainwater, L., Schmaus, G. and Smeeding, T. (1988). Equivalence Scales, Well-Being, Inequality and Poverty: Sensitivity Estimates across Ten Countries Using the Luxembourg Incomes Study Database, *Review of Income and Wealth* **34**: 115-142.

Caminada, K. and Goudswaard, K. (2001). Does a Flat Rate Individual Income Tax Reduce Tax Progressivity? A Simulation for the Netherlands, *Public Finance and Management* **1(4)**: 471-499.

Chakravarty, S.R. (1983). A New Index of Poverty, *Mathematical Social Sciences* **6**: 307-313.

Chakravarty, S.R. (1988). Extended Gini Indices of Inequality, *International Economic Review* **29**: 147-156.

Chakravarty, S.R. and Muliere, P. (2004a). Welfare Indicators: A Review and New Perspectives - 1. Measurement of Inequality, *Metron-International Journal of Statistics* **LXI**(3): 457-497.

Chakravarty, S.R. and Muliere, P. (2004b). Welfare Indicators: A Review and New Perspectives - 2. Measurement of Poverty, *Metron-International Journal of Statistics* **LXII**(2): 247-281.

Champernowne, D.G. (1952). The Graduation of Income Distributions, *Econometrica* **20**: 591-615.

Champernowne, D.G. (1974). A Comparison of Measures of Inequality of Income Distribution, *Economic Journal* **84**: 787-816.

Clark, S., Hemming, R., Ulph, D. (1981). On Indices for the Measurement of Poverty, *Economic Journal* **91**: 515-526.

Coulter, F., Cowell, F.A., Jenkins, S.P. (1992). Equivalence Scale Relativities and the Extent of Inequality and Poverty, *Economic Journal* **102**: 1067-1082.

Cowell, F.A. (1995). Measuring Inequality, 2nd edition, Prentice Hall/Harvester Wheatsheaf, Hemel Hempstead.

Cowell, F.A. (2000). Measurement of Inequality, in Atkinson, A.B. and Bourguignon, F. (2000). Handbook of Income Distribution, pp. 87-166, North-Holland, Amsterdam.

Cowell, F.A. and Kuga, K. (1981). Additivity and the Entropy Concept: An Axiomatic Approach to Inequality Measurement, *Journal of Economic Theory* **25**: 131-143.

Creedy, J. and Sleeman, C. (2005). Adult Equivalence Scales, Inequality and Poverty, Research Paper No. 938, Department of Economics at University of Melbourne, Melbourne.

Dalton, H. (1920). The Measurement of Inequality of Incomes, *Economic Journal* **30**: 348-361.

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**(1): 43-64.

Deutscher Bundestag (1998). 14. Wahlperiode: Jahresgutachten 1998/99 des Sachverstaendigenrates zur Begutachtung der gesamtwirtschaftlichen Entwicklung, Bundestagsdrucksache 14/73.

Deutscher Bundestag (2005). 15. Wahlperiode: Lebenslagen in Deutschland - 2. Armuts- und Reichtumsbericht der Bundesregierung, Bundestagsdrucksache 15/5015.

Duclos, J.-Y. and Araar, A. (2006). Poverty and Equity: Measurement, Policy and Estimation with DAD, Kluwer, Boston/Dordrecht/London.

Duclos, J.-Y., Esteban, J. and Ray, D. (2004). Polarization: Concepts, Measurement, Estimation, *Econometrica* **72**(6): 1737-1772.

Esteban, J., Gradín, C. and Ray, D. (1999). Extensions of a Measure of Polarization with an Application to the Income Distribution of five OECD Countries, Working Paper No. 218, Maxwell School of Citizenship and Public Affairs, Syracuse University, Syracuse, New York.

Esteban, J. and Ray, D. (1994). On the Measurement of Polarization, *Econometrica* **62**(4): 819-851.

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

Foster, J.E. and Shorrocks, A.F. (1988). Poverty Orderings, *Econometrica* **56**: 173-177.

Foster, J.E. and Shorrocks, A.F. (1991). Subgroup Consistent Poverty Indices, *Econometrica* **59**: 687-709.

Gini, C. (1914). Di una Misura della Dissomiglianza tra due Gruppi di quantita e delle sue Applicazioni allo Studio delle Relazioni Statistiche. *Atti R. I. Veneto* **74**: 185-312.

Gradín, C. (2000). Polarization by Sub-Populations in Spain, 1973-1991, *Review of Income and Wealth* **46**(4): 457-474.

Gupta, A. and Kapur, V. (2000). Microsimulation in Government Policy and Forecasting, North-Holland, Amsterdam.

Hagenaars, A.J.M. (1987). A Class of Poverty Indices, *International Economic Review* **28**: 583-607.

Hamada, K. and Takayama, N. (1977). Censored Income Distribution and Measurement of Poverty, *Bulletin of the International Statistical Institute* **47**: 617-632.

Hauser, R. (1999). Personelle Primaer- und Sekundaerverteilung der Einkommen unter dem Einfluss sich aendernder wirtschaftlicher und sozialpolitischer Rahmenbedingungen. Eine empirische Analyse auf der Basis der Einkommens- und Verbrauchsstichproben 1973-1993, *Allgemeines Statistisches Archiv* **83**: 88-110.

Kakwani, N.C. (1977). Measurement of Tax Progressivity: An International Comparison, *Economic Journal* **87**: 71-80.

Kakwani, N.C. (1980). On a Class of Poverty Measures, *Econometrica* **48**: 437-446.

Kakwani, N.C. and Son, H. (2003). On Measures of Inequality and Poverty with Welfare Implications, in A. Shah (ed.), Bringing Civility in Governance, Handbook of Public Sector Performance Reviews, vol. 3, the World Bank, Washington D.C.

Kiefer, D.W. (1985). Distributional Tax Progressivity Measurement, *National Tax Journal* **37**: 497-513.

Kolm, S.-C. (1976). Unequal Inequalities II, *Journal of Economic Theory* **13**: 82-111.

Lambert, P.J. (1994). Measuring Progressivity with Differences in Tax Treatment, in J. Creedy (ed.), Taxation, Poverty and Income Distribution, pp. 17-27.

Lambert, P.J. (2001). The Distribution and Redistribution of Income, 3rd edition, Manchester University Press, Manchester and New York.

Lorenz, M.C. (1905). Methods of Measuring the Concentration of Wealth, *Publications of the American Statistical Association* **9**: 209-219.

Luethi, A. (1981). Messung wirtschaftlicher Ungleichheit, Lecture Notes in Economic and Mathematical Systems No. 189, Springer, Berlin/Heidelberg/New York.

Maasoumi, E. (1997). Empirical Analyses of Inequality and Welfare, in M.H. Pesaran and P. Schmidt (eds.), Handbook of Applied Econometrics, vol. II: Microeconomics, pp. 202-245, Basil Blackwell, Oxford.

Medeiros, M. (2006). Poverty, inequality and redistribution: A methodlgy to define the rich, United Nations Development Programme, Interantional Poverty Center, Working Paper no. 18, Brasilia.

Mehran, F. (1976). Linear Measures of Income Inequality, *Econometrica* **44**: 805-809.

Merz, J., Hirschel, D. and Zwick, M. (2005). Struktur und Verteilung hoher Einkommen – Mikroanalysen auf der Basis der Einkommensteuerstatistik - Beitrag zum zweiten Armuts- und Reichtumsbericht 2004 der Bundesregierung, Bundesministerium für Gesundheit und soziale Sicherung Berlin, (http://www.bmas.bund.de/BMAS/Redaktion/Pdf/Publikationen/struktur-und-vereilung-hoher-einkommen,property=pdf,bereich=bmas,sprache=de,rwb =true.pdf).

Musgrave, R.A. and Thin, T. (1948). Progressive Taxation in an Inflationary Economy, *Journal of Political Economy* **56**: 498-514.

Newson, R. (2004). Generalized Power Calculations for Generalized Linear Models and More, *The Stata Journal* 2004 **4**(4): 379-401.

Nygård, F. and Sandström, A. (1981). Measuring Income Inequality, Almqvist and Wiksell, Stockhom.

Organization for Economic Cooperation and Development (2006). What are Equivalence Scales?, Directorate for Employment, Labour and Social Affairs, (http://www.oecd.org/dataoecd/61/52/35411111.pdf).

Orcutt, G. (1957): A New Type of Socio-Economic Systems, *Review of Economics and Statistics* **58**: 773-797.

Orcutt, G., Merz, J. and Quinke, H. (1986): Microanalytic Simulation Models to Support Social and Financial Policy, North-Holland, Amsterdam.

Peichl, A. (2005). Die Evaluation von Steuerreformen durch Simulationsmodelle, Finanzwissenschaftliche Diskussionsbeitraege Nr. 05-01, Finanzwissenschaftliches Forschungsinstitut an der Universitaet zu Koeln, Cologne.

Peichl, A, Schaefer, T. and Scheicher, C. (2006). Poverty and Richness: Effects of Proposed Tax Reforms for Germany, mimeo, Finanzwissenschaftliches Forschungsinstitut an der Universitaet zu Koeln, Cologne.

Pfaehler, W. (1983). Measuring Redistributional Effects of Tax Progressivity by Lorenz Curves, *Jahrbuecher fuer Nationaloekonomie und Statistik* **198**: 237-249.

Piesch, W. (1975). Statistische Konzentrationsmaße, Mohr, Tuebingen.

Reynolds, M. and Smolensky, E. (1977). Public Expenditures, Taxes, and the Distribution of Income: The United States, 1950, 1961, 1970, Academic Press, New York.

Rodríguez, J.G. and Salas, R. (2002). Extended Bi-Polarization and Inequality Measures. Mimeo, Universidad Complutense de Madrid, Madrid.

Sachverstaendigenrat zur Begutachtung der gesamtwirtschaftlichen Entwicklung (2003). Aktualisierte Analyse der personellen Einkommensverteilung in Deutschland, in Jahresgutachten 2002/03: Zwanzig Punkte fuer Beschaeftigung und Wachstum, Metzler-Poeschel, Stuttgart.

Sachverstaendigenrat zur Begutachtung der gesamtwirtschaftlichen Entwicklung (2004). Verteilung der Markteinkommen und der Einkommensteuerschuld in Deutschland: Eine Auswertung anhand von einkommensteuerlichen Veranlagungsdaten, in Jahresgutachten 2003/04: Staatsfinanzen konsolidieren - Steuersystem reformieren, Metzler-Poeschel, Stuttgart.

Sachverstaendigenrat zur Begutachtung der gesamtwirtschaftlichen Entwicklung (2005). Aktualisierte Analyse der personellen Einkommensverteilung in Deutschland, in Jahresgutachten 2004/05: Erfolge im Ausland, Herausforderungen im Inland, Wiesbaden.

Schmid, F (1993). A General Class of Poverty Measures, *Statistical Papers* **34**: 189-211.

Schmid, F. and Trede, M. (1999). Skript zur Vorlesung Statistische Analyse von Ungleichheit, Armut und Konzentration, Universitaet zu Koeln, Cologne.

Schmidt, A. (2004). Statistische Messung der Einkommenspolarisation, Reihe: Quantitative Oekonomie, Band 141, Eul-Verlag, Lohmar.

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

Shorrocks, A.F. (1980). The Class of Additively Decomposable Inequality Measures, *Econometrica* **48**: 613-625.

Shorrocks, A.F. (1982). Inequality Decomposition by Factor Components, *Econometrica* **50**: 193-211.

Shorrocks, A.F. (1983). Ranking Incomes Distributions, *Economica* **50**: 3-17.

Suits, D. (1977). Measurement of Tax Progressivity, *American Economic Review* **67**: 747-752.

Theil, H. (1967). Economics and Information Theory, North Holland, Amsterdam.

Thon, D. (1979). On Measuring Poverty, *Review of Income and Wealth* **25**: 429-440.

Thon, D. (1983). A Poverty Measure, *Indian Economic Journal* **30**: 55-70.

Vaughan, R. (1987). Welfare Approaches to Measurement of Poverty, *Economic Journal* **97**: 160-170.

Wang, Y.-Q. and Tsui, K.-Y. (2000). Polarization Orderings and New Classes of Polarization Indices, *Journal of Public Economic Theory* **2**(3): 349-363.

Watts, H. (1968). An Economic Definition of Poverty, in D.P. Moynihan (ed.), On Understanding Poverty: Perspectives from the Social Sciences, Basic Books, New York.

Wolfson, M.C. (1994). When Inequalities Diverge, *American Economic Review* **84**(2): 353-358.

Wolfson, M.C. (1997). Divergent Inequalities: Theory and Empirical Results, *Review of Income and Wealth* **43**(4): 401-421.

FINANZWISSENSCHAFTLICHE

DISKUSSIONSBEITRÄGE

Die Beiträge ab 1998 (z.T. auch ältere) stehen auch als kostenloser Download (pdf) zur Verfügung unter: http://www.fifo-koeln.de

1993

93-1 *Ewringmann, D.*: Ökologische Steuerreform? *Vergriffen*

93-2 *Gawel, E.*: Bundesergänzungszuweisungen als Instrument eines rationalen Finanzausgleichs. *Vergriffen*

93-3 *Ewringmann, D. / Gawel, E. / Hansmeyer, K.-H.*: Die Abwasserabgabe vor der vierten Novelle: Abschied vom gewässergütepolitischen Lenkungs- und Anreizinstrument, 2. Aufl. *Vergriffen*

93-4 *Gawel, E.*: Neuere Entwicklungen der Umweltökonomik. *Vergriffen*

93-5 *Gawel, E.*: Marktliche und außermarktliche Allokation in staatlich regulierten Umweltmedien: Das Problem der Primärallokation durch Recht. *Vergriffen*

1994

94-1 *Gawel, E.*: Staatliche Steuerung durch Umweltverwaltungsrecht — eine ökonomische Analyse. *Vergriffen*

94-2 *Gawel, E.*: Zur Neuen Politischen Ökonomie der Umweltabgabe. *Vergriffen*

94-3 *Bizer, K. / Scholl, R.:* Der Beitrag der Indirekteinleiterabgabe zur Reinhaltung von Klärschlamm. *Vergriffen*

94-4 *Bizer, K.:* Flächenbesteuerung mit ökologischen Lenkungswirkungen. *Vergriffen*

1995

95-1 *Scholl, R.*: Verhaltensanreize der Abwasserabgabe: eine Untersuchung der Tarifstruktur der Abwasserabgabe. ISBN 3-923342-39-X. 6,50 EUR

95-2 | *Kitterer, W:* Intergenerative Belastungsrechnungen („Generational Accounting") - Ein Maßstab für die Belastung zukünftiger Generationen? ISBN 3-923342-40-3. 7,50 EUR

1996

96-1 | *Ewringmann, D. / Linscheidt, B. / Truger, A.:* Nationale Energiebesteuerung : Ausgestaltung und Aufkommensverwendung. ISBN 3-923342-41-1. 10,00 EUR

96-2 | *Ewringmann, D. / Scholl. R.:* Zur fünften Novellierung der Abwasserabgabe; Meßlösung und sonst nichts? ISBN 3-923342-42-1. 7,50 EUR

1997

97-1 | *Braun, St. / Kambeck, R.:* Reform der Einkommensteuer. Neugestaltung des Steuertarifs. ISBN 3-923342-43-8. 7,50 EUR

97-2 | *Linscheidt, B. / Linnemann, L.:* Wirkungen einer ökologischen Steuerreform – eine vergleichende Analyse der Modellsimulationen von DIW und RWI. ISBN 3-923342-44-6. 5,00 EUR

97-3 | *Bizer, K. / Joeris, D.:* Bodenrichtwerte als Bemessungsgrundlage für eine reformierte Grundsteuer. ISBN 3-923342-45-4, 7,50 EUR

1998

98-1 | *Kitterer, W.:* Langfristige Wirkungen öffentlicher Investitionen - theoretische und empirische Aspekte. ISBN 3-923342-46-2. 6,00 EUR

98-2 | *Rhee, P.-W.:* Fiskale Illusion und Glory Seeking am Beispiel Koreas (1960-1987). ISBN 3-923342-47-0. 5,00 EUR

98-3 | *Bizer, K.:* A land use tax: greening the property tax system. ISBN 3-923342-48-9. 5,00 EUR

2000

00-1 | *Thöne, M.:* Ein Selbstbehalt im Länderfinanzausgleich?. ISBN 3-923342-49-7. 6,00 EUR

00-2 | *Braun, S., Kitterer, W.:* Umwelt-, Beschäftigungs- und Wohlfahrtswirkungen einer ökologischen Steuerreform : eine dynamische Simulationsanalyse unter besonderer Berücksichtigung der Anpassungsprozesse im Übergang. ISBN 3-923342-50-0. 7,50 EUR

2002

02-1 | *Kitterer, W.:* Die Ausgestaltung der Mittelzuweisungen im Solidarpakt II. ISBN 3-923342-51-9. 5,00 EUR

2005

05-1 | *Peichl, A.:* Die Evaluation von Steuerreformen durch Simulationsmodelle ISBN 3-923342-52-7. 8,00 EUR

05-2 | *Heilmann, S.:* Abgaben- und Mengenlösungen im Klimaschutz : die Interaktion von europäischem Emissionshandel und deutscher Ökosteuer. ISBN 3-923342-53-5. 8,00 EUR

05-3 | *Fuest, C., Peichl, A., Schaefer, T.:* Dokumentation FiFoSiM: Integriertes Steuer-Transfer-Mikrosimulations- und CGE-Modell. ISBN 3-923342-54-3. 8,00 EUR

06-8 | *Bergs, C, Fuest, C., Peichl, A., Schaefer, T.:* Reformoptionen der Familienbesteuerung : Aufkommens-, Verteilungs- und Arbeitsangebotseffekte. ISBN 3-923342-62-4. 6,00 EUR

2006

06-1 | *Fuest, C., Peichl, A., Schaefer, T.:* Führt Steuervereinfachung zu einer „gerechteren" Einkommensverteilung? Eine empirische Analyse für Deutschland. ISBN 3-923342-55-1. 6,00 EUR.

06-2 | *Bergs, C., Peichl, A.:* Numerische Gleichgewichtsmodelle - Grundlagen und Anwendungsgebiete. ISBN 3-923342-56-X. 6,00 EUR.

06-3 | *Thöne, M.:* Eine neue Grundsteuer – Nur Anhängsel der Gemeindesteuerreform? ISBN 3-923342-57-8. 6,00 EUR.

06-4 | *Mackscheidt, K.:* Über die Leistungskurve und die Besoldungsentwicklung im Laufe des Lebens. ISBN 3-923342-58-6. 6,00 EUR

06-5 | *Fuest, C., Peichl, A., Schaefer, T.:* Does tax simplification yield more equity and efficiency? An empirical analysis for Germany. ISBN 3-923342-59-4. 6,00 EUR

06-6 | *Fuest, C., Peichl, A., Schaefer, T.:* Die Flat Tax: Wer gewinnt? Wer verliert? Eine empirische Analyse für Deutschland. ISBN 3-923342-60-8. 6,00 EUR

06-7 | *Kitterer, W.,* Finken, J.: Zur Nachhaltigkeit der Länderhaushalte – eine empirische Analyse. ISBN 3-923342-61-6. 6,00 EUR