A straightforward and versatile calculus for income inequality

updated version

Thomas Kämpke
Forschungsinstitut für anwendungsorientierte Wissensverarbeitung/n FAW/n
Lise-Meitner Str. 9, 89081 Ulm, Germany
kaempke@faw-neu-ulm.de

Abstract

A calculus based on one-parametric Lorenz curves is shown to enable versatile computations over income distributions. These computations focus on empirical as well as on conceptual issues. One empirical issue is the computation of the so-called equity parameter from support points of any Lorenz curve. One conceptual issue is a merger computation that allows for a novel separation of a population into a rich and a poor constituent. This separation allows that the two constituents mix rather than, conventionally, be divided by a given threshold value.

Key words: Index numbers, Lorenz curves, welfare aggregation.

1 Introduction

Quantitative analysis of income distribution has regained attention in the last years in the context of sustainable pathways into the future. In the following, the definition of poverty by the European Union where an individual is considered to be poor when falling short of 50% of the average income of the whole population, is shown to lead to a differential equation as well as to an equivalent functional equation on conditional expectations. The solutions of the differential equation allow to deal with a variety of empirical as well as systems-theoretical issues. These issues do not refer to poverty alone but to the whole income distribution and they represent the abstract core of work which has been pursued over some years [KPR].

The solutions of the underlying differential equation form a one-parametric class of Lorenz curves. This allows to compute quantiles, cumulative and average incomes of population segments, minimum income levels and the like. All these will be expressed as a function or computing scheme of the one parameter which is denoted as equity parameter. One or the other of the present issues could be addressed by other approaches. But uniformity is achieved by a calculus which is based on the equity parameter. This is pivotal to the present report.

All computations are intended to be simple and straightforward in order to allow for closing the gap between conceptual and empirical issues. With the exceptions of regressions and population merger, the computations can be executed in closed form. Whenever this is not the case, numerical approximations must be resorted to. The computation of the joint Lorenz curve for the merger of two populations each having its own Lorenz curve is facilitated by a convolution-like operation.

The remainder of this paper is organized as follows. A certain type of one-parametric Lorenz curves is derived in a concrete and in a more abstract way in sections 2 and 3 respectively. This leads to the so-called equity parameter which serves as an inequality index of the income distribution. Equity parameters are computed by curve fitting to empirical data in section 4. A relation to another index for income distributions, the Atkinson index, is established in section 5. The celebrated issue of whether any relation might hold between growth and income distribution is touched in section 6. The evenly important issue of an optimal level of distributional inequality is investigated from a middle class perspective in section 7. Section 8 deals with a somewhat speculative comparison of high income groups from nations with either very uneven or more balanced income distributions.

Section 9 assumes that a society is split into two subsocieties each having its own income distribution. The two distributions are typically interleaved which means that each accounts for incomes that are higher than some incomes from the other distribution. Such situations result from the merger or from alliance-building of two nations. The perspective taken here is that such a fusion has occured and that a split into two subsocieties is considered in retrospect. This kind of subsociety analysis is distinct from any split of income groups according to a given income threshold.

The similarity between the subsociety distributions in terms of average income and equity parameter as well as between the size of each subsociety describes the homogeneity of the overall income distribution. The findings of this approach are the conceptually and computationally most difficult results of the proposed calculus. But following the empirical results only (section 9.2) allows to skip their derivation (section 9.1). Moreover, the sections after section 2 are independent to a great extent.

2 Differential equation

2.1 Approach

The cumulative distribution of income or consumption of a nation is described by a Lorenz curve $F(x)$, $x \in [0,1]$. The cumulative income is related to individual incomes by the derivative of the Lorenz curve. This can be seen by considering an individual at income rank x who is one with $100 \cdot x\%$ of the population earning less and 100% - $100 \cdot x\%$ of the population earning more than him. This individual adds $F'(x)$ to the values of the Lorenz curve whenever the curve is absolutely continuous meaning that $F(x) = \int_0^x F'(u)du$; the latter is assumed throughout. Unless otherwise stated, the total income of a nation is normalized to unity which means that $F(1) = 1$.

The foregoing consideration is related to the relativity concept of poverty of the European Union. According to this notion, an individual of some nation is poor if his income falls short of 50% of the average per capita income of that nation [EU], [FI]. Instead of considering only the poorest, any individual's income will be compared to the average of all larger incomes. Moreover, the actual fraction of individual vs. average income need not be 50% but some other, yet unknown value.

Comparing individual incomes to average incomes leads to a differential equation [KPR]. The rationale is as follows. The segment $[0, x]$ of the population receives its proportion $F(x)$ of income. The remaining income $1-F(x)$ is distributed among the remaining fraction of the population which is $1-x$. The average income of all richer individuals thus equals $\frac{1-F(x)}{1-x}$. An individual at income rank x is supposed to have a constant fraction ε, $\varepsilon < 1$, thereof. Combining this with the previous fact that an individual contribution to the Lorenz curve is nothing but the curve's derivative, results in the linear inhomogenous differential equation

$$F'(x) = \varepsilon \cdot \frac{1 - F(x)}{1 - x}.$$

All solutions of this differential equation that satisfy the normalization conditions $F(0) = 0$ and $F(1) = 1$ are given by the manifold of Pareto distributions $F_\varepsilon(x) = 1 - (1 - x)^\varepsilon$. The parameter ε whose values range between zero and one is called equity parameter.

The density of a Pareto Lorenz curve is $f_\varepsilon(x) = F'_\varepsilon(x) = \varepsilon\,(1 - x)^{\varepsilon-1}$ so that $f_\varepsilon(0) = \varepsilon$. The intuition of the last equation is that the theoretically smallest of all incomes is exactly at the level of the equity parameter. In absolute terms this means that the theoretically smallest income equals the average per capita income multiplied by the equity parameter.

Distributional inequality of Pareto Lorenz curves decreases in the equity parameter both in the sense of Lorenz dominance and Gini index. The first means that Lorenz curves of the Pareto type do not intersect for different parameters except at the endpoints and the curve with smaller parameter lies below the curve with larger parameter. The second means that the Gini index of Pareto Lorenz curves is decreasing in the equity parameter

$$\text{Gini-Index} = 2 \cdot \int_0^1 x - F_\varepsilon(x)dx = \frac{1 - \varepsilon}{1 + \varepsilon}.$$

As an alternative to Lorenz curves, income distributions can be described by histograms or random variables which either denote absolute incomes or multiples of the average income. The latter means that value one amounts to the average income, value two amounts to twice the average income etc. Let W denote such a random variable. The foregoing differential equation then transforms to a functional equation of conditional expectations. This functional equation is

$$w = \varepsilon \cdot E(W| \, W \geq w)$$

for all $w \in [\varepsilon, \infty)$. In strict analogy to the differential equation, the equity parameter appears as multiple of an expression of all incomes that lie above some income level which is specified by the left side of the equation. Whenever the random variable has a density function, the functional equation obviously transforms to the integral equation

$$w = \varepsilon \cdot \frac{\int_w^\infty \, u\varphi(u)\, du}{\int_w^\infty \, \varphi(u)\, du}$$

for all $w \in [\varepsilon, \infty)$.

2.2 Related work

The differential equation approach can be variated to lead to other one-parametric as well as to two-parametric Lorenz curves where the equity parameter is replaced by a so-called equity function. The current type of Lorenz curve and the more general form $F(x) = (1-(1-x)^\alpha)^{\frac{1}{\beta}}$, $0 < \alpha < 1$, $\beta \leq 1$ has also been investigated by Rasche et al. [Ra]. This type of curve apparently was motivated by curvature features. Driven by empirical motivation, β-distributions $F(x) = x - \vartheta x^\gamma \, (1-x)^\delta$ as well as quadratic income distributions $F(x) = \frac{1}{2}(bx + e + \sqrt{mx^2 + nx + e^2})$ were also considered [Da]. An overview on parametric Lorenz curves is given in [ChGr] and parametric Lorenz curves that were merely adopted from probability distributions are treated in [RySl].

Modifications of the Rasche curves such as the Lorenz curves $F(x) = x^\alpha(1-(1-x)^\beta)$, $\alpha > 0$, $0 < \beta \leq 1$, and exponential curves $F(x) = \frac{e^{\kappa \, x}-1}{e^\kappa - 1}$, $\kappa > 0$, have also been proposed, see [Che]. In addition to parametric Lorenz curves, non-parametric approaches such as kernel estimators and quantile ratios [Gi] have been proposed.

An interesting class of inequality indices is formed by measures that require other than income or consumption data like entropy measures and the Atkinson index [Lit]. The latter needs an external parameter. A relation between this parameter and the equity parameter will be established below. A survey of inequality measurement is given in [Sil].

A recent network approach [BouMez] for finite many economical agents and a similar systems dynamics approach [NiSo] both claim that the heavy tail of absolute wealth is approximately Pareto-distributed. This result from "econophysics" is coherent with the present assumptions which can be reformulated as the income distribution adhering to a so-called power law. Here, the validity of the power law is assumed throughout the whole distribution.

3 Self-similarity

A completely different derivation of the presently investigated type of Lorenz curves stems from self-similarity. Though it may seem intuitively obvious, the formal concept of self-similarity for income analysis is less clear. In particular, any relation to self-similarity as used in non-linear dynamics, in particular in so-called chaos theory, is believed to be misleading here. The reason is that non-linear dynamics considers more complicated phenomena related to iteration schemes or partition processes as for space-filling curves and the like.

Self-similarity, today, is a popular concept that appeals to scientific imagination but may severly lack clarity. In particular, self-similarity can be vaguely identified with or be very clearly related to other issues like power laws; see [LADW] for a rigorous treatment of the latter. Self-similarity attains its highest potential when being tailored to a specific problem rather than being stylized as a single, universal notion that is probably loaded with some mystics. Here, a particular notion of self-similarity for Lorenz curves is given.

Income distributions are often described by Lorenz curves. Thus, it is a straightforward quest to directly relate self-similarity to Lorenz curves. This relation will be established by truncation and normalization operations.

The Lorenz curve $F(x)$ of a whole population allows to derive the Lorenz curve of any fixed population segment $[x_0, 1]$, called the rich segment, by rescaling the population fraction – which means rescaling the argument – and by normalizing to the cumulative income of the rich segment. This results in the truncated Lorenz curve

$$F^{x_0}(x) = \frac{F(x_0 + x\,(1 - x_0)) - F(x_0)}{1 - F(x_0)}$$

for $0 \leq x \leq 1$.

3.1 Self-similarity by pointwise equality

The original Lorenz curve is understood to be self-similar if all truncated Lorenz curves are equal to the original curve so that

$$F^{x_0}(x) = F(x)$$

for $0 \leq x \leq 1$ and $0 \leq x_0 < 1$. This functional equation for self-similarity is illustrated in figure 1. The

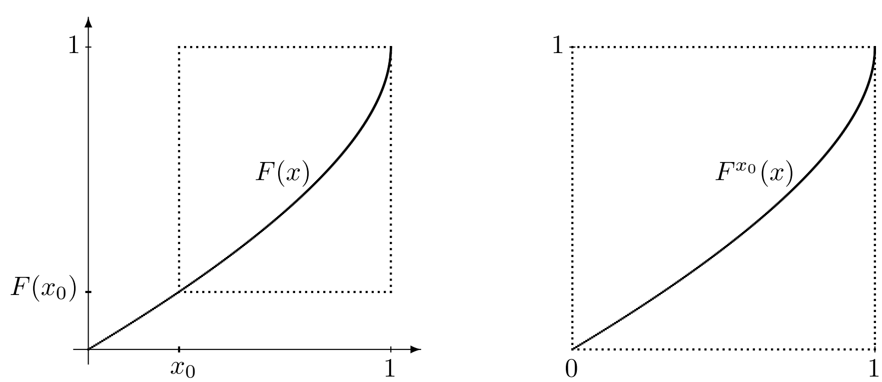

Figure 1: The dashed box and the enclosed section of the Lorenz curve (left) are rescaled to the unit square (right). Self-similarity requires the complete curve $F(x)$ and the curves $F^{x_0}(x)$ to be equal for all truncation values $0 \leq x_0 < 1$.

Pareto Lorenz curves from section 2 are self-similar which is easily verified as follows.

$$\begin{aligned} F_\varepsilon^{x_0}(x) &= \frac{F_\varepsilon(x_0 + x\,(1 - x_0)) - F_\varepsilon(x_0)}{1 - F_\varepsilon(x_0)} \\ &= \frac{1 - (1 - x_0 - x\,(1 - x_0))^\varepsilon - 1 + (1 - x_0)^\varepsilon}{1 - 1 + (1 - x_0)^\varepsilon} \\ &= 1 - (\frac{1 - x_0 - x\,(1 - x_0)}{1 - x_0})^\varepsilon \\ &= 1 - (1 - x)^\varepsilon \quad = \quad F_\varepsilon(x). \end{aligned}$$

More interesting, the converse is also true. Any self-similar Lorenz curve $F(x)$ which is differentiable and has a continuous derivative at zero is of the form $F(x) = 1 - (1-x)^{F'(0)}$. This means that $F(x) = F_\varepsilon(x)$ for equity parameter $\varepsilon = F'(0)$.

The claim can be verified as follows. First, an argument substitution leads to an equivalent functional equation for self-similarity. The substitution is $w = w(x) = x_0 + x\,(1-x_0)$. The new argument ranges between x_0 and 1 for the original argument x ranging between zero and one. The original argument is expressed by the new argument as $x = \frac{w-x_0}{1-x_0}$. Thus, the self-similarity equation becomes

$$\frac{F(w)-F(x_0)}{1-F(x_0)} = F(\frac{w-x_0}{1-x_0}).$$

This allows the subsequent transformation of the functional equation to a differential equation by formally denoting the new argument w again by x.

$$\begin{aligned}
\frac{F(x)-F(x_0)}{1-F(x_0)} = F(\frac{x-x_0}{1-x_0}) \implies & \frac{F(x)-F(x_0)}{x-x_0} = (1-F(x_0)) \cdot \frac{F(\frac{x-x_0}{1-x_0})}{x-x_0} \\
\implies & F'(x_0) = \lim_{x\to x_0} \frac{F(x)-F(x_0)}{x-x_0} = (1-F(x_0)) \cdot \lim_{x\to x_0} \frac{F(\frac{x-x_0}{1-x_0})}{x-x_0} \\
& \overset{0/0}{=} (1-F(x_0)) \cdot \lim_{x\to x_0} \frac{F'(\frac{x-x_0}{1-x_0})}{1} \cdot (\frac{x-x_0}{1-x_0})' \\
& = (1-F(x_0)) \cdot \lim_{x\to x_0} \frac{F'(\frac{x-x_0}{1-x_0})}{1} \cdot \frac{1}{1-x_0} \\
& = \frac{1-F(x_0)}{1-x_0} \cdot F'(0).
\end{aligned}$$

This is the linear inhomogenous differential equation from section 2 for $\varepsilon = F'(0)$ for which the solution manifold was given by the Lorenz curves $F_\varepsilon(x)$.

Interestingly, the functional equation for self-similarity does not require the Lorenz curves to be of any parametric type. But all solutions of the functional equation are parametric showing that the equity parameter evolves from self-similarity.

3.2 Self-similarity by equality of Gini-indices

Pointwise equality of the original Lorenz curve with all truncations can be relaxed to equality of Gini-indices. An income distribution is therefore understood to be self-similar if each truncation has the same Gini-index as the original Lorenz curve which means that

$$2 \cdot \int_0^1 x - F^{x_0}(x)dx = 2 \cdot \int_0^1 x - F(x)dx$$

for all $x_0 \in (0,1)$. The condition is sketched for only one truncation value in figure 2.

Each Lorenz curve which is self-simiar in this sense is of the Pareto type. The argument is as follows. Abbreviating the Gini-index by

$$g = 2 \cdot \int_0^1 x - F(x)dx = 1 - 2\int_0^1 F(x)dx,$$

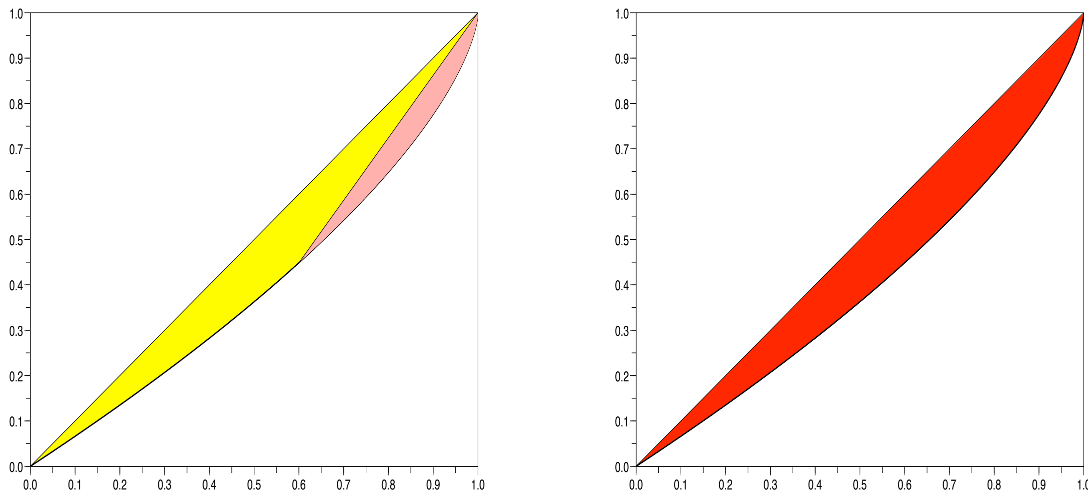

Figure 2: Truncation at $x_0 = 0.6$. The Gini-index of the original curve, which equals the size of the two shaded areas (left) must be equal to the size of the shaded area (right) which is derived from the small shaded area (left) by the same zooming operation as in figure 1.

the stipulated equality of all Gini-indices results in the identity

$$\int_0^1 F^{x_0}(x)dx = \int_0^1 F(x)dx = \frac{1-g}{2}.$$

Thus

$$\begin{aligned}
&\int_0^1 \frac{F(x_0 + x\,(1-x_0)) - F(x_0)}{1-F(x_0)}dx = \frac{1-g}{2} \\
&\implies \int_0^1 F(x_0 + x\,(1-x_0))dx = \frac{1-g}{2}(1-F(x_0)) + F(x_0) \\
&\implies \int_{x_0}^1 \frac{F(w)}{1-x_0}dw = F(x_0)(1-\frac{1-g}{2}) + \frac{1-g}{2} \\
&\implies \int_{x_0}^1 F(w)dw = F(x_0)(1-x_0)\frac{1+g}{2} + \frac{1-g}{2}(1-x_0) \\
&\implies \frac{d}{dx_0}\int_{x_0}^1 F(w)dw = \frac{d}{dx_0}\Big[F(x_0)(1-x_0)\frac{1+g}{2} + \frac{1-g}{2}(1-x_0)\Big] \\
&\implies -F(x_0) = \frac{1+g}{2}\Big[F'(x_0)(1-x_0) - F(x_0)\Big] - \frac{1-g}{2} \\
&\implies 0 = \frac{1+g}{2}(1-x_0)F'(x_0) - F(x_0)(\frac{1+g}{2} - 1) - \frac{1-g}{2}
\end{aligned}$$

The second implication uses the same substitution $w = w(x) = x_0 + x(1-x_0)$ that was used for self-similarity by pointwise equality. This substitution has the differentials $dw/dx = 1-x_0$ or $dx = dw/(1-x_0)$.

The fifth implication uses the fundamental theorem of calculus. All in all, the sequence of equations can be continued as

$$\begin{aligned} F'(x_0) &= \frac{F(x_0)(\frac{1+g}{2}-1)}{\frac{1+g}{2}(1-x_0)} + \frac{1-g}{2(1-x_0)\frac{1+g}{2}} \\ &= \frac{1}{1-x_0}\left[\frac{1-g}{1+g} + F(x_0)\frac{1+g-2}{1+g}\right] \\ &= \frac{1-g}{1+g}\cdot\frac{1-F(x_0)}{1-x_0}. \end{aligned}$$

Thus, each Lorenz curve which is self-similar in the sense of Gini-indices satisfies the differential equation from section 2 with $\varepsilon = \frac{1-g}{1+g}$ and, thus, is of the type $F_\varepsilon(x)$.

4 Empirics

The Lorenz curves of type F_ε can be fitted to empirical data from support sets in the usual sense of sum of least squares. The set support set is denoted as $\{(x_i, y_i)|\, i = 1, \ldots, n\}$. This amounts to a solution of the regression problem

$$\min_{0<\varepsilon<1} \sum_{i=1}^{n} (F_\varepsilon(x_i) - y_i)^2.$$

Thus regression problem is not known to be solvable in closed form so that approximations must be resorted to. The fitting objective is not a convex function of the equity parameter but it is a quasiconvex function. This still means that it has a unique local minimum which is global. The minimizer is denoted as best fit equity parameter.

For approximation, parameter sweeping was applied here which means that the error function was evaluated for a finite number of candidate parameters and the minimzer of these was selected as best fit equity parameter.

Computations of regressions were performed on the so-called world development indicators of the World Bank [Wo]. The support points for Lorenz curves of 30 nations as well as their best fit equity parameters are given in the subsequent table.

Nation	x_1	y_1	x_2	y_2	x_3	y_3	x_4	y_4	x_5	y_5	x_6	y_6	ε
Austria	0.1	0.044	0.2	0.104	0.4	0.252	0.6	0.437	0.8	0.666	0.9	0.807	0.6494
Brazil	0.1	0.009	0.2	0.025	0.4	0.08	0.6	0.18	0.8	0.363	0.9	0.524	0.2778
Canada	0.1	0.028	0.2	0.075	0.4	0.204	0.6	0.376	0.8	0.606	0.9	0.762	0.5525
China	0.1	0.022	0.2	0.055	0.4	0.153	0.6	0.302	0.8	0.525	0.9	0.691	0.4464
Czech Rep.	0.1	0.043	0.2	0.103	0.4	0.248	0.6	0.425	0.8	0.642	0.9	0.776	0.6173
Denmark	0.1	0.036	0.2	0.096	0.4	0.245	0.6	0.428	0.8	0.655	0.9	0.795	0.6289
Finland	0.1	0.042	0.2	0.1	0.4	0.242	0.6	0.418	0.8	0.641	0.9	0.784	0.6135
France	0.1	0.028	0.2	0.072	0.4	0.198	0.6	0.37	0.8	0.598	0.9	0.749	0.5376
Germany	0.1	0.037	0.2	0.09	0.4	0.225	0.6	0.4	0.8	0.629	0.9	0.774	0.5882
Gr. Britain	0.1	0.026	0.2	0.066	0.4	0.181	0.6	0.344	0.8	0.571	0.9	0.727	0.5025
Greece	0.1	0.03	0.2	0.075	0.4	0.199	0.6	0.368	0.8	0.596	0.9	0.747	0.5376
Hungary	0.1	0.039	0.2	0.088	0.4	0.213	0.6	0.379	0.8	0.602	0.9	0.752	0.5525
India	0.1	0.035	0.2	0.081	0.4	0.197	0.6	0.347	0.8	0.54	0.9	0.665	0.4673
Italy	0.1	0.035	0.2	0.087	0.4	0.227	0.6	0.408	0.8	0.637	0.9	0.782	0.5988
Japan	0.1	0.048	0.2	0.106	0.4	0.248	0.6	0.424	0.8	0.644	0.9	0.783	0.6211
Korean Rep.	0.1	0.029	0.2	0.075	0.4	0.204	0.6	0.378	0.8	0.607	0.9	0.757	0.5525
Mexico	0.1	0.014	0.2	0.036	0.4	0.108	0.6	0.226	0.8	0.418	0.9	0.572	0.3279
Netherlands	0.1	0.028	0.2	0.073	0.4	0.2	0.6	0.372	0.8	0.6	0.9	0.749	0.5405
Nigeria	0.1	0.016	0.2	0.044	0.4	0.126	0.6	0.251	0.8	0.444	0.9	0.592	0.3546
Norway	0.1	0.041	0.2	0.1	0.4	0.243	0.6	0.422	0.8	0.646	0.9	0.788	0.6211
Poland	0.1	0.03	0.2	0.077	0.4	0.203	0.6	0.37	0.8	0.591	0.9	0.737	0.5319
Portugal	0.1	0.031	0.2	0.073	0.4	0.189	0.6	0.348	0.8	0.566	0.9	0.716	0.5000
Russian Fed.	0.1	0.017	0.2	0.044	0.4	0.13	0.6	0.263	0.8	0.464	0.9	0.613	0.3731
S. Africa	0.1	0.011	0.2	0.029	0.4	0.084	0.6	0.176	0.8	0.353	0.9	0.541	0.2809
Slovakia	0.1	0.051	0.2	0.119	0.4	0.277	0.6	0.463	0.8	0.685	0.9	0.818	0.6806
Spain	0.1	0.028	0.2	0.075	0.4	0.201	0.6	0.371	0.8	0.598	0.9	0.748	0.5405
Sweden	0.1	0.037	0.2	0.096	0.4	0.241	0.6	0.422	0.8	0.654	0.9	0.799	0.6289
Switzerland	0.1	0.026	0.2	0.069	0.4	0.196	0.6	0.369	0.8	0.598	0.9	0.748	0.5376
USA	0.1	0.015	0.2	0.048	0.4	0.153	0.6	0.313	0.8	0.548	0.9	0.715	0.4673
Venezuela	0.1	0.013	0.2	0.037	0.4	0.121	0.6	0.257	0.8	0.469	0.9	0.63	0.3788

As a reminder for the interpretation of the equity parameter, a value of $\varepsilon = 0.5376$ (France) means that each income approximately equals 53.76% of the average of all higher incomes. The interpretation is ever more appropriate when the regression error becomes ever smaller. The largest regression error of the previous best fits is attained by USA and the smallest best fit is attained by the Czech Republic as well as by Slovakia. All best fit regression errors are listed in [KPR].

5 Relation to the Atkinson index

Out of the plethora of indices for income inequality, the Atkinson index [A] is one that not merely aggregates an income distribution to a single number but which requires a so-called shape parameter. The effect of this parameter is to make the measure more sensitive either to incomes that lie close to or far from the average income. The Atkinson index is one of the "newer" measures of income inequality which actually is considered in official statistics, see for example [USCensus].

For a finite discrete distribution of absolute incomes $w_1, \ldots, w_n$, the Atkinson index is defined as

$$A_e = 1 - \Big[\frac{1}{n}\sum_{i=1}^{n} \big(\frac{w_i}{\bar{w}}\big)^{1-e}\Big]^{1/(1-e)}.$$

All absolute incomes are transformed to relative incomes by normalizing with the average income $\bar{w} = 1/n \cdot \sum_{i=1}^{n} w_i$. The shape parameter e varies between 0 and 1. This parameter is arbitrarily and externally set; in particular, it does not depend on empirical data. Increasing shape parameters lead to increasing values of the Atkinson index starting from zero since $A_e = 0$ for arbitrary distributions and for $e = 0$. This is expressed as "... the higher this parameter the more society is concerned about inequality" [Lit] though this intuition is not easy to grasp. The Atkinson index itself has the same indication as the Gini index. Large values denote large inequality and small values denote little inequality for any fixed positive shape parameter.

It is proposed here to use equity parameters as shape parameters for the Atkinson index. Therefore, best fit equity parameters are computed by regression as in section 4 and then inserted into the formula. This allows to compare different income distributions by the Atkinson index with a certain range of shape parameters. For illustration, two hypothetical distributions of six absolute incomes are considered.

w_i	10	11	12	20	21	22
v_i	10	11	12	30	31	32

Both distributions are quite even, since the largest income is only the 2.2 fold and – respectively – the 3.2 fold of the smallest income. For three arbitrarily chosen values of the shape parameter the values of the Atkinson index are computed as follows.

	$e = 0.25$	$e = 0.50$	$e = 0.75$
w	0.01285	0.02582	0.03881
v	0.02273	0.06091	0.09169

What is a reasonable value or interval of values for the shape parameter when comparing the two distributions? The best fit equity parameter for the two distributions is obtained from normalizing the cumulative income to one. This results in the support points and the best fit equity parameter of the two Lorenz curves as follows.

	1/6	2/6	3/6	4/6	5/6	Best fit equity parameter
w	10/96	21/96	33/96	53/96	74/96	$\varepsilon = 0.7095$
v	10/126	21/126	33/126	63/126	94/126	$\varepsilon = 0.6071$

Computing the Atkinson index for the first, more even distribution with shape parameter $e = 0.7095$ and for the second, less even distribution with shape parameter $e = 0.6071$ leads to comparisons by the subsequent range of values.

	$e = 0.6071$	...	$e = 0.7095$
w	0.03138	...	0.03671
v	0.07413	...	0.08673

The Atkinson indices are thus computed only for certain values above one half in this example. The previously mentioned concern about inequality is supposed to be actually expressed in terms of equity parameters. It is hence recommended to use the range of these parameters for the shape parameters.

6 Growth and equity revisited

The most promising potential to reduce inequality is assumed to originate from growth which is expressed by the view that "... economic growth is an indispensable requirement for poverty reduction" [UK]. Growth itself is not formally modelled here in order to not become susceptible to disputes about growth theories as such. The present view of growth is pragmatic and closer to exogenous than to endogenous concepts as it calls for transfers without a formally rational decision like utility maximization or else by an economical agent.

A quantitative relation between growth and inequality or between growth and another indicator has often been searched for. A celebrated example of such a hypothetical relation is given by Kuznets' inverse-U shaped curves. This hypothesis states that the average income of a nation and, thus, growth as a function of inequality is inverse-U shaped [GaKu]. Inverse-U shaped curves typically are parabolas with negative curvature parameter. However, various data analyses have shown that there is little if no correlation between growth and other parameters when the search is carried out "blindly" across nations. Some studies including [Mi] indicate that the income distribution on a global scale is recently becoming more uneven as growth is very different in different parts of the world.

Growth and inequality are considered as being determined by common factors rather than being independent variables [LuSq], [PR]. Then, starting out from an unequal income distribution, growth and inequality reduction may go hand in hand. An example is Singapur.

Historic data of Singapur's economic development show quite a strong growth at an average annual rate of about 4.5% after inflation over the years from 1980 to 1994. For the same period, the income distribution becomes more even [SI]. The given data on deciles of absolute income levels transform to the subsequent support points for Lorenz curves.

1980		1994	
x_i	y_i	x_i	y_i
0.1	0.007738	0.1	0.01696
0.2	0.03608	0.2	0.04694
0.3	0.07486	0.3	0.08739
0.4	0.12349	0.4	0.13914
0.5	0.18434	0.5	0.20295
0.6	0.26042	0.6	0.28239
0.7	0.35606	0.7	0.38134
0.8	0.47996	0.8	0.50660
0.9	0.65013	0.9	0.67372

The strong growth goes along with a development process that reduces inequality; low income groups benefit more from the economic development than high income groups. This may be considered as so-called capacity building. Ever larger population segments receive better education, better health care, better opportunities to play an increasingly important role in the economy etc. The devlopment is reflected by an increase in the best fit equity parameter from $\varepsilon_{1980} = 0.3803$ to $\varepsilon_{1994} = 0.4206$.

The foregoing example does not suggest or even imply that certain levels of the equity parameter lead to substantial growth rates nor is the converse true. Rather, for certain regions, growth and inequality may turn out to be endogenously related. This is a matter of the regime in control.

7 Towards optimal levels of inequality

The issue of an optimal or ideal level of inequality for income distributions has often been raised either in relation to growth, see above, or in relation to welfare. For the latter see for example [Gr], [Ko]. The intention here is to elaborate on the welfare aspect from a middle class perspective.

In accordance with the underlying notion of poverty which relates the lowest income to the average of all incomes, middle class incomes are also defined in terms of the average of all incomes. This can be achieved by expressing middle class incomes by multiples of the lowest theoretical income ε; the average of all incomes is again normalized to unity. Middle class incomes are understood as intervals $[\beta\varepsilon, \alpha\varepsilon]$ for multiples $\alpha > \beta > 1$.

Since the middle class can be considered as the work horse of a prosperous society, its cumulative income should be maximized among varying equity parameters. Formally, this amounts to the solution of the following maximization problem.

$$\max_{0<\varepsilon<1} F_\varepsilon(x_{\alpha\varepsilon}) - F_\varepsilon(x_{\beta\varepsilon}),$$

where the arguments of the Lorenz curve are the unique points given by the respective density equations $f_\varepsilon(x_{\alpha\varepsilon}) = \alpha\varepsilon$ and $f_\varepsilon(x_{\beta\varepsilon}) = \beta\varepsilon$. These density equations can be solved explicitly by $x_{\alpha\varepsilon} = 1 - exp(\frac{\ln\alpha}{\varepsilon-1})$ with the same formula for β. A critical point analysis then shows that the maximizer for the middle class can be computed in closed form as

$$\varepsilon = \frac{\ln\frac{\ln\alpha}{\ln\beta}}{\ln\frac{\ln\alpha}{\ln\beta} + \ln\frac{\alpha}{\beta}}.$$

Sample values of formally optimal equity parameters are as follows.

Middle class income boundaries		Optimal equity parameter	Middle class range	Middle class size
			for optimal equity parameter	
β	α	ε	$[x_{\beta\varepsilon}, x_{\alpha\varepsilon}]$	$x_{\alpha\varepsilon} - x_{\beta\varepsilon}$
2	5	0.4790	$[0.7356, 0.9544]$	0.2188
1.5	5	0.5338	$[0.5809, 0.9683]$	0.3874
2	10	0.4272	$[0.7018, 0.9820]$	0.2802
1.1	10	0.5906	$[0.2077, 0.9964]$	0.7887
1.01	100	0.5718	$[0.0230, 0.9999]$	0.9769

Optimal distributional inequality is at equity parameter levels slightly above one half in many situations including the extreme case that is specified in the last line of the foregoing table. That situation assigns almost 98% of the whole population to the middle class. Defining the middle class to have at least the double of the theoretically lowest income appears to leave out segments that intuitively belong to the middle class. Thus, the lower boundary should be less than two. Then, the best fit equity parameters of many European and some Asian nations as computed in section 4 are close to the optimal values.

8 Comparing absolute income levels

A developed nation typically does not only have a higher per capita income than an undeveloped nation but, also, has a more even income distribution. Often, the spread between the absolute income distributions is so large, that most absolute incomes in the undeveloped nation are smaller than in the developed nation. But there need not be dominance meaning that this relation need not hold for all incomes. Exceptions and even a reversal may apply to very high incomes.

Since empirical data on very high incomes are difficult to obtain for one nation [Pa] and even more so for a comparison between nations, we attain a purely formal point of view. The Pareto Lorenz curves are assumed to denote exactly the income distributions at the heavy tail. Then it turns out that there is a critical muiltiple of the per capita income of the developed nation so that the population fraction with at least this absolute income is larger in the undeveloped nation than in the developed nation.

Formally, let $\alpha(n, \varepsilon, \varepsilon')$ be the critical multiple as a function of the equity parameter ε of the developed nation, the equity parameter ε' of the undeveloped nation and the multiple n of the per capita income of the developed nation vs. the undeveloped nation. The critical level can be computed in closed form by

$$\alpha(n, \varepsilon, \varepsilon') = \varepsilon \cdot (n \frac{\varepsilon}{\varepsilon'})^{\frac{1-\varepsilon}{\varepsilon-\varepsilon'}}.$$

Sample values of the critical multiple are given in the following table for a per capita ratio of $n = 5$.

ε' undeveloped nation	ε developed nation							
	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.1	20 mio	3921.4	160.00	27.951	9.1169	4.1413	2.2952	1.4484
0.2	-	400452	400.00	33.663	9.0000	3.8987	2.1715	1.4041
0.3	-	-	35117	100.23	12.927	4.4189	2.2546	1.4134
0.4	-	-	-	4768.4	33.750	6.1250	2.5298	1.4604
0.5	-	-	-	-	777.60	12.964	3.2000	1.5588
0.6	-	-	-	-	-	138.95	5.3333	1.7617
0.7	-	-	-	-	-	-	26.122	2.2819
0.8	-	-	-	-	-	-	-	5.0625
minimum value	1.737 mio	3347.9	159.99	27.144	8.6024	3.8774	2.1715	1.4028
ε'_0	0.0501	0.0751	0.1001	0.1252	0.1502	0.1753	0.2003	0.2253

The entries of the diagonal section of this table are read as follows. When comparing a developed nation with $\varepsilon = 0.6$ and an undeveloped nation with $\varepsilon' = 0.3$ whose ratio of per capita incomes is $n = 5$, then the population segment in the undeveloped nation having an absolute income of at least the 12.927-fold of the per capita income of the developed nation is larger than in the developed nation. In this sense there are more high incomes in undeveloped nations than in developed nations.

The entries from the diagonal section of the foregoing table are not monotone when considered columnwise. The minimum equity parameter for the undeveloped nation is columnwise denoted by ε'_0. Formally, this is the minimizer of the critical income levels so that $\varepsilon'_0 = argmin_{0<\varepsilon'<\varepsilon}\alpha(n, \varepsilon, \varepsilon')$. This minimizer does not admit a closed form solution but must be approximated numerically. The corresponding critical value is specified as "minimum value" in the second last row of the table. This is the least unfavourable critical value for an undeveloped nation in comparison to a fixed equity parameter of the developed nation and a fixed ratio of per capita incomes between the two nations.

9 Inspection of homogeneity by merger analysis

Whenever two distinct societies are known by their characteristics which here are populations (size), per capita incomes and Lorenz curves, they can be merged to form one society with joint Lorenz curve. The merger may be real as for nations forming a union or it may be conceptual. The intricate computation of the common Lorenz curve is derived below.

The inverse merger problem can be stated for a single society. The society is assumed to be split into two subsocieties as in an Apartheid situation. However, low and high incomes may be present in both subsocieties. This makes the subsequent decomposition of a society distinct from decompositions by thresholding such as into rich and poor. The issue at hand is to identify both subsocieties only from empirical data. This will be facilitated by a multivariate regression. The characteristics of the two subsocieties form the variables of the regression.

Whenever the equity parameters and the per capita incomes of the two subsocieties are close, the whole society is understood to be homogenous. Whenever the equity parameters or the per capita incomes are far, the whole society is understood to be inhomogenous. The assignment of an individual to either of the subsocieties is impossible since the approach is based on holistic data.

The empirical results of the merger analysis (section 9.2) can be followed without going through the intricate derivation (subsections 9.1.1 and 9.1.2).

9.1 Merger

It can be expected for the joint Lorenz curve that it bears more inequality than any of the individual Lorenz curves. The reason is that individual populations may have similar income distributions on clearly distinct absolute levels. Merging the populations may then cause more unevenness; equivalently, the inter-inequality dominates the intra-inequality. This means that the equity parameter of the joint Lorenz curve or a parametric approximation therof, need not lie between the equity parameters of the individual Lorenz curves. But when income levels are close, the equity parameter of the joint Lorenz curve may actually lie between the individual equity parameters as in the case of India, see below.

The formal situation is described as follows.

Characteristics	Subsociety 1	Subsociety 2
Equity parameter	ε_1	ε_2
Lorenz curve	$F_{\varepsilon_1}(x) = 1 - (1-x)^{\varepsilon_1}$	$F_{\varepsilon_2}(x) = 1 - (1-x)^{\varepsilon_2}$
Density	$f_{\varepsilon_1}(x) = \varepsilon_1\,(1-x)^{\varepsilon_1 - 1}$	$f_{\varepsilon_2}(x) = \varepsilon_2\,(1-x)^{\varepsilon_2 - 1}$
Population	m	$1-m$
cumulative income	b	$1-b$

The difficulty of computing the joint Lorenz curve is that the subsocieties mix at different rates when arranged according to increasing absolute income. An additional difficulty stems from the different cases that determine the merger and the non-merger sectors of the joint Lorenz curve. The two sectors are illustrated in figure 3. In the lower sector, the poorest of one population exclusively make up the poorest of the merged population. Then proper merging occurs. Since the densities in figure 3 indicate absolute

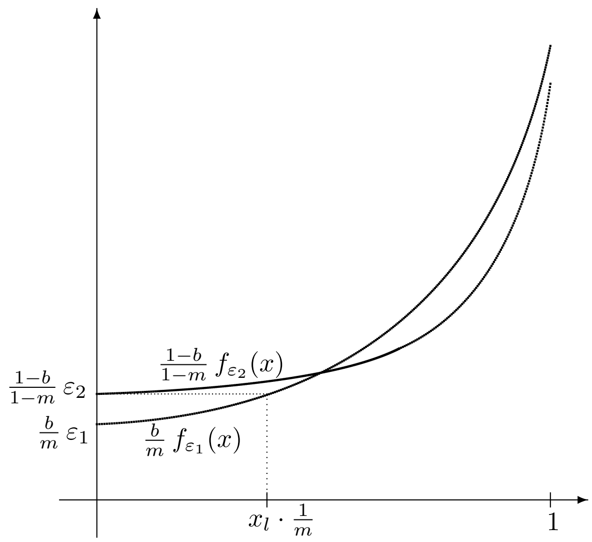

Figure 3: Two densities of per capita income which are given as products of the average incomes with the derivatives of Lorenz curves. The merger of the two subsocieties begins when income values attain common values. The domain of these income levels is referred to as merger area.

incomes, the integrals below the curves add up to b/m and $1 - b/(1-m)$ respectively rather than unity. Both densities are unbounded. As a consequence, the merger area reaches right to the upper interval boundary; there is no high income sector which is made up of one population only.

9.1.1 Merger formulas

The joint Lorenz curve depends on the four parameters $\varepsilon_1, \varepsilon_2, m, b$. All these range between zero and one. The joint Lorenz curve is given as a mixture of the individual Lorenz curves

$$F_{\varepsilon_1,\varepsilon_2,m,b}(x) = b\, F_{\varepsilon_1}(x_1\, \frac{1}{m}) + (1-b)\, F_{\varepsilon_2}(x_2\, \frac{1}{1-m}),$$

for those arguments x that have a deconvolution $x = x_1 + x_2$ or, equivalently, $x_2 = x - x_1$ which is understood to be the unique solution x_1 of the deconvolution equation

$$\frac{b}{m} f_{\varepsilon_1}(x_1 \frac{1}{m}) = \frac{1-b}{1-m} f_{\varepsilon_2}((x - x_1)\frac{1}{1-m}).$$

The solution lies in the open interval $(\max\{0, x-1+m\}, \min\{x, m\})$. This interval can be used as search region for any numerical procedure that actually computes the deconvolution. The deconvolution cannot be computed in closed form except for trivial parameter constellations.

For arguments that do not have a deconvolution, the joint Lorenz curve is given by

$$F_{\varepsilon_1,\varepsilon_2,m,b}(x) = \begin{cases} (1-b)\, F_{\varepsilon_2}(\frac{1}{1-m}\, x) & \text{if } (1-b)/(1-m)\cdot\varepsilon_2 \le b/m\cdot\varepsilon_1 \\ b\, F_{\varepsilon_1}(\frac{1}{m}\, x) & \text{if } (1-b)/(1-m)\cdot\varepsilon_2 > b/m\cdot\varepsilon_1 \end{cases}$$

for $0 \le x \le x_l$. The value x_l is the lower bound of the merger area. This value is specified for the joint Lorenz curve so that it receives a normalizing factor when visualized in terms of the individual income distributions as in figure 3. The lower bound of the merger area lies between zero and one and it is given as the unique solution of the equations

$$\begin{cases} (1-b)/(1-m)\cdot f_{\varepsilon_2}(\frac{1}{1-m}\, x) = b/m\cdot\varepsilon_1 & \text{for } (1-b)/(1-m)\cdot\varepsilon_2 \le b/m\cdot\varepsilon_1 \\ b/m\cdot f_{\varepsilon_1}(\frac{1}{m}\, x) = (1-b)/(1-m)\cdot\varepsilon_2 & \text{for } (1-b)/(1-m)\cdot\varepsilon_2 > b/m\cdot\varepsilon_1. \end{cases}$$

The lower bound of the merger area can be computed in closed form. For the two foregoing cases specified in the foregoing order the values are

$$x_l = \begin{cases} (1-m)\cdot\Big(1-exp(\frac{\ln b-\ln(1-b)+\ln(1-m)-\ln m+\ln\varepsilon_1-\ln\varepsilon_2}{\varepsilon_2-1})\Big) \\ m\cdot\Big(1-exp(\frac{\ln(1-b)-\ln b+\ln m-\ln(1-m)+\ln\varepsilon_2-\ln\varepsilon_1}{\varepsilon_1-1})\Big). \end{cases}$$

The merger area $(x_l, 1]$ is exactly that range of population quantiles x that have a deconvolution.

9.1.2 Derivation

For the sake of the derivation of the joint Lorenz curve, let the subsocieties consist of M_i individuals with cumulative income B_i, $i = 1, 2$. The income at rank x is then typically formed by i_0 individuals from the first and j_0 individuals from the second subsociety so that

$$x \approx \frac{i_0 + j_0}{M_1 + M_2} = \underbrace{\frac{i_0}{M_1 + M_2}}_{=x_1} + \underbrace{\frac{j_0}{M_1 + M_2}}_{=x_2}.$$

Both individual incomes at rank x are equal up to discretization error and both individuals are at different ranks within their subsocieties which requires rescaling of the argument. This is formally expressed as

$$\frac{B_1}{M_1} f_{\varepsilon_1}(\frac{i_0}{M_1}) \approx \frac{B_2}{M_2} f_{\varepsilon_2}(\frac{j_0}{M_2}).$$

This obviously results in

$$\frac{B_1}{M_1} f_{\varepsilon_1}(\frac{i_0}{M_1+M_2}\frac{M_1+M_2}{M_1}) \approx \frac{B_2}{M_2} f_{\varepsilon_2}(\frac{j_0}{M_1+M_2}\frac{M_1+M_2}{M_2}).$$

Ignoring the discretization error results in the equation

$$\frac{B_1}{M_1} f_{\varepsilon_1}(x_1\frac{M_1+M_2}{M_1}) = \frac{B_2}{M_2} f_{\varepsilon_2}(x_2\frac{M_1+M_2}{M_2}).$$

Inserting the original parameters $B_1 = b, M_1 = m, B_2 = 1-b, M_2 = 1-m$ results in the deconvolution equation

$$\frac{b}{m} f_{\varepsilon_1}(x_1\frac{1}{m}) = \frac{1-b}{1-m} f_{\varepsilon_2}(x_2\frac{1}{1-m}).$$

All in all, the income at rank x of the joint distribution equals

$$\begin{aligned}
&\frac{B_1}{M_1} f_{\varepsilon_1}(0) + \frac{B_1}{M_1} f_{\varepsilon_1}(\frac{1}{M_1}) + \ldots + \frac{B_1}{M_1} f_{\varepsilon_1}(\frac{i_0-1}{M_1}) \\
&\qquad + \frac{B_2}{M_2} f_{\varepsilon_2}(0) + \frac{B_2}{M_2} f_{\varepsilon_2}(\frac{1}{M_2}) + \ldots + \frac{B_2}{M_2} f_{\varepsilon_2}(\frac{j_0-1}{M_2}) \\
&= \sum_{i=1}^{i_0} \frac{1}{M_1}\, B_1 f_{\varepsilon_1}(\frac{i-1}{M_1}) + \sum_{j=1}^{j_0} \frac{1}{M_2}\, B_2 f_{\varepsilon_2}(\frac{j-1}{M_2}) \\
&\approx \int_0^{\frac{i_0}{M_1}} B_1 f_{\varepsilon_1}(u)du + \int_0^{\frac{j_0}{M_2}} B_2 f_{\varepsilon_2}(u)du \\
&= \int_0^{\frac{i_0}{M_1+M_2}\frac{M_1+M_2}{M_1}} B_1 f_{\varepsilon_1}(u)du + \int_0^{\frac{j_0}{M_1+M_2}\frac{M_1+M_2}{M_2}} B_2 f_{\varepsilon_2}(u)du \\
&\approx B_1 F_{\varepsilon_1}(x_1\frac{M_1+M_2}{M_1}) + B_2 F_{\varepsilon_2}(x_2\frac{M_1+M_2}{M_2}).
\end{aligned}$$

The sums are Riemann sums of the given integrals and thus provide approximations. Normalizing from absolute to relative per capita income requires to divide by the total income $B_1 + B_2 = 1$. This leads to the Lorenz curve

$$F_{\varepsilon_1,\varepsilon_2,m,b}(x) = b\, F_{\varepsilon_1}(x_1\, \frac{1}{m}) + (1-b)\, F_{\varepsilon_2}(x_2\, \frac{1}{1-m}),$$

for any quantile x that is uniquely decomposable into $x = x_1 + x_2$ with solution x_1 of the deconvolution equation. The formula for the Lorenz curve on the lower sector, where no deconvolution exists, is derived similarly.

9.2 Regression

The joint Lorenz curves can be fitted to empirical data from support sets $\{(x_i, y_i) | \, i = 1, \ldots, n\}$ in the same way as the individual Lorenz curves, see section 4. This amounts to a solution of the regression problem

$$\min_{0 \le \varepsilon_1,\varepsilon_2,m,b \le 1} \sum_{i=1}^{n} (F_{\varepsilon_1,\varepsilon_2,m,b}(x_i) - y_i)^2.$$

This regression problem is multivariate with four regression parameters: the equity parameters of the two subsocieties $\varepsilon_1, \varepsilon_2$ and the relative size m and the relative income share b of one subsociety. The relative size and the relative income share of the other subsociety are then simply computed as $1-m$ and $1-b$.

The multivariate regression can approximately be solved by enumeration at fixed step sizes of the four independent variables. Enumeration at adaptive step sizes and even gradient procedures are alternative methods. The computations of the joint Lorenz curves can be performed in closed form with one exception. The deconvolution over the merger area (see above) is obtained by numerical enumeration. Thus, the whole regression consists of nested numerical procedures which pile up to a surprisingly high computational load.

9.2.1 Nations

The support data of nations for the regression are again the world development indicators that were used for the one-parametric regression in section 4. The regression results are stated in the following table[1]. The poorer subsociety is listed first for all nations which means that the average incomes in the poorer subsociety, in the overall society and in the richer subsociety are arranged as $\frac{b}{m} < 1 < \frac{1-b}{1-m}$.

[1]The computations for this table and for the world income (see below) were done in MATLAB by Markus Stark.

Nation	ε_1	b	m	b/m	ε_2	$1-b$	$1-m$	$(1-b)/(1-m)$	ε
Austria	0.92	0.16	0.30	0.53	0.72	0.84	0.70	1.20	0.6494
Brazil	0.52	0.16	0.56	0.29	0.38	0.84	0.44	1.81	0.2778
Canada	0.78	0.12	0.30	0.40	0.64	0.88	0.70	1.26	0.5535
China	0.48	0.36	0.60	0.60	0.60	0.64	0.40	1.60	0.4464
Czech Rep.	0.64	0.28	0.40	0.70	0.68	0.72	0.60	1.20	0.6173
Denmark	0.86	0.14	0.28	0.50	0.70	0.86	0.72	1.19	0.6289
Finland	0.76	0.18	0.32	0.56	0.68	0.82	0.68	1.21	0.6135
France	0.74	0.12	0.30	0.40	0.62	0.88	0.70	1.26	0.5376
Germany	0.86	0.14	0.30	0.47	0.66	0.86	0.70	1.23	0.5882
Gr. Britain	0.88	0.10	0.30	0.33	0.58	0.90	0.70	1.29	0.5025
Greece	0.86	0.14	0.34	0.41	0.62	0.86	0.66	1.30	0.5376
Hungary	0.58	0.48	0.64	0.75	0.70	0.52	0.36	1.44	0.5525
India	0.38	0.52	0.56	0.93	0.62	0.48	0.44	1.09	0.4673
Italy	0.82	0.14	0.30	0.47	0.68	0.86	0.70	1.23	0.5988
Japan	0.82	0.20	0.34	0.59	0.68	0.80	0.66	1.21	0.6211
Korean Rep.	0.72	0.14	0.32	0.44	0.64	0.86	0.68	1.26	0.5525
Mexico	0.40	0.28	0.58	0.48	0.42	0.72	0.42	1.71	0.3279
Netherlands	0.78	0.12	0.30	0.40	0.62	0.88	0.70	1.26	0.5405
Nigeria	0.78	0.08	0.32	0.25	0.40	0.92	0.68	1.35	0.3546
Norway	0.92	0.14	0.28	0.50	0.68	0.86	0.72	1.19	0.6211
Poland	0.78	0.14	0.32	0.44	0.60	0.86	0.68	1.26	0.5319
Portugal	0.84	0.14	0.34	0.41	0.56	0.86	0.66	1.30	0.5000
Russ. Fed.	0.62	0.12	0.38	0.32	0.44	0.88	0.62	1.42	0.3731
S. Africa	0.22	0.66	0.82	0.80	0.94	0.34	0.18	1.89	0.2809
Slovakia	0.84	0.22	0.34	0.65	0.76	0.78	0.66	1.18	0.6806
Spain	0.72	0.14	0.32	0.44	0.62	0.86	0.68	1.26	0.5405
Sweden	0.82	0.18	0.34	0.53	0.72	0.82	0.66	1.24	0.6289
Switzerland	0.68	0.16	0.36	0.44	0.64	0.84	0.64	1.31	0.5376
USA	0.64	0.10	0.34	0.29	0.58	0.90	0.66	1.36	0.4673
Venezuela	0.80	0.06	0.32	0.19	0.46	0.94	0.68	1.38	0.3788

The table is best read starting with the data for South Africa. The size data of the subsocieties roughly match the nation's former Apartheid situation; the richer subsociety consists of about 18% of the population and the poorer of about 82%. Remarkably, this result follows from only one Lorenz curve, namely that of the whole nation.

Assuming a bipartition for each other nation leads to subsocieties that are more balanced in size though the income inequality may even be larger as in the comparison of Brazil with South Africa. An explanation is that the income ranks between the very top and the very bottom in all other nations are more densly occupied than in South Africa – irrespective of the degree of inequality.

It is no surprise that the overall equity parameter lies below the equity parameters of each of the two subsocieties in almost all cases. The only exceptions are India and South Africa. The reason for the more frequent case is that the difference between the average income levels creates extra inter-inequality in addition to the intra-inequality within each subsociety, see above. But for South Africa the equity parameters are very far apart and for India the two subsocieties have very similar income levels (0.93 and 1.09) so that inter-inequality is almost negligible. Another observation about India's subsocieties is that they differ clearly from the US subsocieties though both nations have equal equity parameters.

In all nations except Brazil, China, Hungary, India, Mexico and South Africa the richer subsociety is larger than the poorer subsociety. Brazil and India have equally sized subsocieties with Mexico's subsocieties being very similar-sized to these.

With the exceptions of China and the Czech Republic the larger of the two subsocieties has the smaller of the two equity parameters. One reason for the more frequent situation of the larger subsociety having the smaller equity parameter is that, all other things being equal, the larger subsociety tends to have a larger income variation as the smaller subsociety.

9.2.2 World

The world as a whole can be analyzed in the same way as individual nations. The support data for the Lorenz curve of the World-income were computed in a "one man one income" mode. This means that income distributions were weighted by population size of each nation under consideration. Worldbank data from more than 150 nations were aggregated and some data gaps had to be filled-in.

The resulting income distribution is more uneven than that of any single nation. The world equity parameter equals 0.12 [StK]. The quantile data are as follows.

"Nation"	x_1	y_1	x_2	y_2	x_3	y_3	x_4	y_4	x_5	y_5	x_6	y_6
World	0.1	0.0028	0.2	0.009	0.4	0.023	0.6	0.0501	0.8	0.1261	0.9	0.3133

The four-parameter regression results in the subsequent subsociety data.

"Nation"	ε_1	b	m	b/m	ε_2	$1-b$	$1-m$	$(1-b)/(1-m)$	ε
World	0.06	0.64	0.82	0.78	0.90	0.36	0.18	2.00	0.12

According to this computation, the world population can be assumed to consist of two subsocieties which have the same proportions as those of South Africa. But the world's spread in terms of equity parameters and average incomes is even more polarized.

References

[A] Atkinson, A.B., "On the measurement of inequality", Journal of Economic Theory 2, 1970, p. 244-263.

[BouMez] Bouchard, J.-P., Mezard, M., "Wealth condensation in a simple model of economy", Physica A, 282, 2000, p. 536ff.

[Che] Cheong, K.S., "An empirical comparison of alternative functional forms for the Lorenz curve", Applied Economics Letters 9, 2002, p. 171-176; www2.soc.hawaii.edu/ econ/workingpapers/992.pdf.

[ChGr] Chotikapanich, D., Griffiths, W.E., "Estimating Lorenz curves using a Dirichlet Distribution", Working paper 110, University of New England, Armidale, 1999; www.une.edu.au/febl/EconStud/EMETwp110.pdf.

[Da] Datt, G., "Computational tools for poverty measurement and analysis", discussion paper 50, International Food Policy Research Institute, Washington, 1998; www.ifpri.cgiar.org/divs/fcnd/dp/dp50.htm.

[EU] European Parliament, "The fight against poverty in the ACP countries and in the European Union", Working Document, Strassbourg, 1999; www.europarl.eu.int/dg2/acp/stras99/en/genrep9.htm#ANNEX I.

[FI] Ministry of Social Affairs and Health, Finland, internet presentation; www.stm.fi/english/tao/publicat/poverty/definit.htm.

[GaKu] Galbraith, J.K., Kum, H., "Inequality and economic growth: data comparisons and econometric tests", LBJ School of Public Affairs, working paper #21, University of Texas, Austin, 2002; utip.gov.utexas.edu/web/workingpaper/UTIP21rv.pdf.

[Gi] Ginther, D.K., "A nonparametric analysis of the US earnings distributions", Discussion paper 1067-95, University of Wisconsin, Madison, 1995; www.ssc.wisc.edu/irp/pubs/dp106795.pdf.

[Gr] Grodner, A., "Optimal wealth inequality in the economy with heterogenous workers and interdependent preferences", report, University of Syracuse, Syracuse, 2003.

[K] Kämpke, T., "Functional inequality measurement", Empirical Economics 29, 2004, p. 697-704.

[KPR] Kämpke, T., Pestel, R., Radermacher, F.J., "A computational concept for normative equality", European Journal of Law and Economics 15, 2003, p. 129-163.

[Ko] Koch, J.V. (ed.), "The distribution of household incomes in Hampton Roads", report, Old Dominion University, 2003; www.odu.edu/bpa/forecasting/index.shtml.

[LADW] Li, L., Alderson, D., Doyle, J.C., Willinger, W., "Towards a theory of scale-free graphs: definition, properties, and implications", Internet Mathematics 2, 2005, p. 431-523, www.internetmathematics.org.

[Lit] Litchfield, J.A., "Inequality: methods and tools", Discussion paper, Worldbank, Washington, 1999.

[LuSq] Lundberg, M., Squire, L., "The simultaneous evolution of growth and inequality", The Economic Journal 113, 2003, p. 326-344.

[Mi] Milanovic, B., "True world income distribution, 1988 and 1993: first calculation based on household surveys alone", The Economic Journal 112, 2002, p. 51-92.

[NiSo] Nirei, M., Souma, W., "Income distribution and stochastic multiplicative process with reset events", in: The Complex Dynamics of Economic Interaction, Springer, New York, 2003.

[Pa] Parenti, M., "The super rich are out of sight", Political Archive, internet presentation, 2000; www.michaelparenti.org/Superrich.html.

[PR] Pestel, R., Radermacher, F.J., "Equity, wealth and growth", Manuscript to the EU Project TERRA 2000, FAW, Ulm, 2003.

[Ra] Rasche, R.H., Gaffney, J., Koo, A.Y.C., Obst, N., "Functional forms for estimating the Lorenz curve", Econometrica 48, 1980, p. 1061-1062.

[RySl] Ryu, H.K., Slotje, D.J., "Parametric approximations of the Lorenz curve", in [Sil], p 291-312.

[Sil] Silber, J. (ed.), "Handbook of income inequality measurement", Kluwer, Boston, 1999.

[SI] Singapur, "Income growth and distribution", report, Dept. of Statistics, Singapur, 1995; www.singstat.gov.sg/papers/op/op-s3.pdf.

[StK] Stark, M., Kämpke, T., "The world income distribution and its equality", FAW technical report #02004, Ulm, 2002, 14 p.

[UK] UK Secretary of State for International Development, "Eliminating world poverty: making globalization work for the poor", white paper, 2000; www.globalisation.gov.uk.

[USCensus] USCensus, "The changing shape of the nation's income distribution", report, US Census Bureau, US Dept. of Commerce, 2000.

[Wo] Worldbank, "World Development Report 2000/2001", Washington, 2001.