Equality functions

Thomas Kämpke and Markus Stark
Forschungsinstitut für anwendungsorientierte Wissensverarbeitung FAW
Helmholtzstr. 16, 89081 Ulm, Germany
{kaempke,stark}@faw.uni-ulm.de

Abstract

The equality parameter as a highly aggregated index of an income distribution is complemented by the finer measure of an equality function. Lorenz curves with linear equality functions are developed and fitted to empirical data by squared error regression. The transition from equality parameters to equality functions significantly reduces regression errors. Relative poverty lines are subsequently introduced, computed approximately and their induced rankings of nations are investigated.
Key words: differential equation, income distribution, Lorenz curve.

1 Introduction

Income distributions and poverty measurement continue to receive attention in particular with markets becoming ever more global. The two issues of income distribution and poverty measurement are typically considered separately. The relative poverty notion of the European Union, which considers an individual to be poor when falling short of 50% of the average income of the whole population, allows a certain combined view. This has earlier been shown to lead to the Pareto class of Lorenz curves. This curve type is generalized here by replacing the so-called equality parameter with functions, so-called equality functions. These functions specify values of equality parameters that are allowed to vary with the income level.

The transition from equality parameters to equality functions is facilitated by a differential equation. The original form of this differential equation characterizes the Pareto distribution. Two generalizations thereof are considered here which one being given in explicit and the other in implicit form.

The equality parameter is some sort of average over an equality function. The Lorenz curves belonging to equality functions exhibit a significantly better fit to empirical data than the Lorenz curves belonging to equality parameters. Also, Lorenz curves belonging to equality functions allow to reasonably operate with relative poverty lines. Unless absolute poverty lines such as "fraction of population with per capita income of at most US $1 per day", relative poverty lines indicate the fraction of the population whose income falls short of 50% of the overall average income. Relative poverty lines are numerically computable.

The parameters of the Lorenz curves will be computed by regression in the usual sense of minimum squared error fits. Since the regressions here are not computable in closed form, approximate fits are computed by an enumeration scheme. It will turn out that continuous Pareto distributions lead to best fits in most cases.

This work is organized as follows. equality functions and some of their features are formally introduced in section 2. One class of Lorenz curves considered requires complete numerical handling saince it does not admit an explicit formula. This makes regressions for curve fitting computationally cumbersome. These regressions based on data for individual nations and for an aggregated world income are given in section 3. Also, numerical approximations of relative poverty lines are computed. Section 4 draws a short conclusion.

2 Approach

2.1 Equality function

According to the EU poverty definition, those individuals are considered poor whose income falls short of half of the overall average income, comp. [EU], [FI]. When this view is extended from the poorest to all others, i.e. when each income is compared to the average of all larger incomes, and when the proportional constant may differ from 0.5 ("half"), the resulting Lorenz curve $y(x)$ satisfies the differential equation $y'(x) = \varepsilon \frac{1-y(x)}{1-x}$, see [KPR]. The ratio $\frac{1-y(x)}{1-x}$ denotes the average of all incomes of the $(1-x) \cdot 100\%$ richest of the population. Noteworthy, all Lorenz curves are normalized to attain values between zero and one so that the sum of all incomes equals unity. On the scale of the derivative unity amounts to the average value of all incomes. The constant value ε is called the equality parameter which attains values between zero and one.

To provide for more generality we consider the differential equation

$$y'(x) = \varepsilon(x)\,\frac{1-y(x)}{1-x}$$

over the domain $0 \leq x < 1$. The function $\varepsilon(x) : [0,1] \to (0,1]$ is assumed to be continuous over its domain. This function is called equality function. The derivative of each Lorenz curve that satisfies the differential equation is bounded from below by its equality function meaning that $y'(x) \geq \varepsilon(x)$ for all $x \in [0,1)$. This lower bound is always attained at the lower bound of the income scale meaning that $y'(0) = \varepsilon(0)$.

For constant functions ("equality parameters") as well as for discontinuous but intervalwise constant functions the differential equation has a closed form solution. For the constant function $\varepsilon(x) = \varepsilon$ the solution is given by $y_P(x) = 1-(1-x)^{\varepsilon}$. The absolute income distribution belonging to this Lorenz curve is of the Pareto type and, thus, the Lorenz curve itself is here understood to be of the Pareto type. This slight deviation from standard terminology will simplify the subsequent notation.

For intervalwise constant equality functions the solution is intervalwise of the Pareto type with continuous "connections". For more general equality functions, even for linear functions $\varepsilon(x) = \varepsilon_1 + (\varepsilon_2 - \varepsilon_1)\,x$ as in figure 1 the differential equation appears to have no closed form solution.

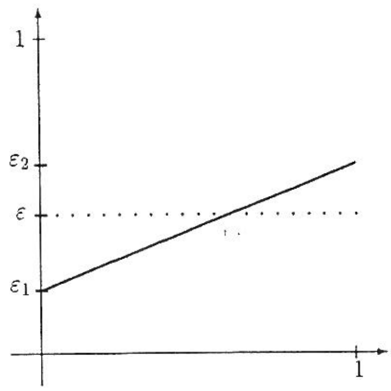

Figure 1: Increasing linear equality function and equality parameter.

The solution of the differential equation is denoted as implicit Pareto distribution $y_{iP}(x)$. Though it generalizes the Pareto distribution, it is different from other forms of generalized Pareto distributions which have different domains and admit closed form representations, comp. [EmKlMi].

The function $y_{cP}(x) = 1 - (1-x)^{\varepsilon(x)}$, called continuous Pareto distribution, is a straightforward generalization of the ordinary Pareto distribution. It satisfies the differential equation

$$y'(x) = \varepsilon(x)\,\frac{1-y(x)}{1-x} - \varepsilon'(x)\,\frac{1-y(x)}{1-x}\cdot(1-x)\ln(1-x).$$

This differential equation includes the derivative of the equality function and other terms and thus is not of an immediate form that allows the equality function to be considered as proportional factor of the average of all larger incomes. Thus, it is considered for approximations. These approximations show a consistent over- or underestimation for monotone equality functions as given in section 2.3.1. For non-monotone equality functions no consistent deviation from the true values is known.

It appears that there is no differential equation with closed form solution that "slightly" generalizes the original case; either the differential equation becomes significantly more complicated (continuous Pareto distribution) or the function itself is more complicated (implicit Pareto distribution). Both the continuous and the implicit Pareto distributions have their individual equality functions which generally differ.

2.2 Related work

The ordinary Pareto-type Lorenz curve has been proposed, among others, by Rasche et al. [Ra], even in the more general form $y(x) = (1-(1-x)^{\alpha})^{\frac{1}{\beta}}$, $0 < \alpha < 1$, $\beta \leq 1$. This type of curve apparently was motivated by insufficient curvature features of other types of Lorenz curves. Driven by empirical motivation, β-distributions $y(x) = x - \vartheta x^{\gamma}(1-x)^{\delta}$ and quadratic income distributions $y(x) = \frac{1}{2}(bx + e + \sqrt{mx^2 + nx + e^2})$ were considered, see [Da]. An overview on parametric Lorenz curves is provided in [ChoGri]. Quite a few parametric Lorenz curves were adopted from probability distribution functions, see for example [RySl, p. 294].

Variations of the Rasche curves like $y(x) = x^{\alpha}(1-(1-x)^{\beta})$, $\alpha > 0$, $0 < \beta \leq 1$, and the exponential curves $y(x) = \frac{e^{\kappa x}-1}{e^{\kappa}-1}$, $\kappa > 0$, have also been proposed, see [Che]. Parametric Lorenz curves are complemented by non-parametric curves such as kernel estimates and by quantile ratios, see [Gi]. The latter is motivated by simplicity issues.

The differential equation leading to the ordinary Pareto distribution can be modified to lead to other one-parametric income distributions. This requires convexity analysis combined with an overconstraining approach [K] and results in Lorenz curves which exhibit a better fit to empirical data than the Pareto distribution [KR]. However, the involved parameters do no longer allow an easy interpretation as equality parameter.

A recent network approach for a finite set of economical agents concludes that the far tail or heavy tail of absolute wealth is approximately Pareto-distributed, see [BouMe]. This result from econophysics is coherent with the present approach.

2.3 Dominance properties

2.3.1 Under- and overestimation

For monotone equality functions the continuous Pareto distribution can be considered from the viewpoint of the wrong differential equation which is the differential equation of the implicit Pareto distribution. The continuous Pareto distribution satisfies the inequalities

$$y'_{cP}(x) \geq \varepsilon(x)\,\frac{1-y_{cP}(x)}{1-x} \text{ resp. } y'_{cP}(x) \leq \varepsilon(x)\,\frac{1-y_{cP}(x)}{1-x}$$

for increasing resp. decreasing equality functions. Assuming that Lorenz curves of the continuous Pareto-type are the true distributions leads to values of an equality function being too small in case the equality function is increasing. This means that when expressing a particular income as a fraction of the average of all larger incomes, this fraction is underestimated.

2.3.2 Lorenz dominance

An important relation between Lorenz curves is given by a partial order resulting from non-intersection. This order, called Lorenz dominance, amounts to the dominating curve lying below the other over the whole domain, comp. [Lit]. Lorenz dominance is nothing but stochastic dominance when Lorenz curves are considered as probability dsitribution functions. The ordinary and the continuous Pareto distribution are related for increasing equality functions by the dominance relations

$$1-(1-x)^{\varepsilon(0)} \leq 1-(1-x)^{\varepsilon(x)} \leq 1-(1-x)^{\varepsilon(1)}.$$

For decreasing equality functions the relations become

$$1-(1-x)^{\varepsilon(1)} \leq 1-(1-x)^{\varepsilon(x)} \leq 1-(1-x)^{\varepsilon(0)}.$$

The continuous Pareto distribution is thus sandwiched by ordinary Pareto distributions in the sense of Lorenz dominance.

Lorenz dominance preserves comparability of equality functions for continuous as well as for implicit Pareto distributions. This means that whenever two equality functions do not intersect, then their Lorenz curves do not intersect. For the continuous Pareto distribution this is formally denoted as the implication

$$\varepsilon_1(x) \leq \varepsilon_2(x) \text{ for all } x \in [0,1] \Rightarrow 1-(1-x)^{\varepsilon_1(x)} \leq 1-(1-x)^{\varepsilon_2(x)} \text{ for all } x \in [0,1].$$

This kind of dominance relation can be demonstrated easily. Comparability of equality functions for the implicit Pareto distribution is more complicated to derive. A numerical approach such as given in section 2.4 allows to deduce the implication

$$\varepsilon_1(x) \leq \varepsilon_2(x) \text{ for all } x \in [0,1] \Rightarrow y_{iP,\varepsilon_1(\,)}(x) \leq y_{iP,\varepsilon_2(\,)}(x) \text{ for all } x \in [0,1].$$

Interestingly, a common equality function implies a dominance relation between the continuous and the implicit Pareto distribution. The relation is given for an increasing resp. decreasing equality function by

$$y_{iP}(x) \leq y_{cP}(x) \text{ resp. } y_{iP}(x) \leq y_{cP}(x).$$

These inequalities are motivated – not formally derived – by comparing the right hand sides of the corresponding differential equations. The intricate deduction of these inequalities is based on uniformly approximating the equality function by monotone step functions and comparing the resulting Lorenz curves in explicit form. This argument is omitted here.

The dominance relation between the continuous and the implicit Pareto distribution has an implication for empirical analysis. This implication is illustrated only for increasing equality functions as this is the prevailing monotonicity direction on real data, see below. Whenever distributions of the two types are fitted to the same data, the implicit Pareto distribution must be lifted in comparison to the continuous distribution when starting out from a common equality function. Lifting the distributions is achieved by lifting the equality function as specified by the foregoing comparability relation. Thus, it can be expected that values of equality functions of the implicit Pareto distribution are larger than those of the continuous Pareto distribution when approximating the same empirical data. This will turn out to be true in section 3.2.

2.4 Numerical solution

Linear equality functions will be considered from now on only. The resulting Lorenz curve is then two-parametric. The parameters are the lower equality parameter ε_1 and the upper equality parameter ε_2. The implicit Lorenz curve can be given approximately by a numerical solution of the discretized ordinary differential equation

$$\frac{y(x_i)-y(x_{i-1})}{x_i-x_{i-1}} = (\varepsilon_1+(\varepsilon_2-\varepsilon_1)\,x_i)\,\frac{1-y(x_i)}{1-x_i},\ i=1,\ldots,n,$$

where $0 = x_0 < x_1 < \ldots < x_n = 1$ and $y(0) = 0$. Approximating the differential at x_i by the symmetric difference ratio $\frac{y(x_{i+1})-y(x_{i-1})}{x_{i+1}-x_{i-1}}$ is infeasible here since already the case $\varepsilon_1 = \varepsilon_2$ leads to non-differentiability at the right interval boundary. Thus, the functional value at the right interval boundary should not go into approximations of the derivative. The differential at x_i is only approximated from the left.

The linear system for Lorenz curve approximation can be solved by forward computation resulting in the values $y_{iP,\varepsilon_1,\varepsilon_2}(x_i)$. The values lie on an increasing and convex function for all feasible parameter values. Thus, the function values can be considered as sample values of a valid Lorenz curve.

3 Empirical findings

3.1 Regression approach

The parameters of a linear equality function are identified by a standard least squares approach

$$\min_{\varepsilon_1,\varepsilon_2} \sum_{x_i \in Z'} (y_{iP,\varepsilon_1,\varepsilon_2}(x_i) - y_i)^2$$

with support set $Z' \subset \{x_0, \ldots, x_n\}$. The value y_i denotes the value of the empirical Lorenz curve given at x_i. The empirical values must be given externally over the whole support set. Both the lower and the upper equality parameter vary over the same, finite candidate set. The candidate set for the two parameters being the same gives enough freedom to allow the best fit equality function to be increasing, decreasing or constant.

The regression approach for the continuous Pareto distribution is given in the analogous form

$$\min_{\varepsilon_1,\varepsilon_2} \sum_{x_i \in Z'} (y_{cP,\varepsilon_1,\varepsilon_2}(x_i) - y_i)^2,$$

with the regression function being explicitly given as $y_{cP,\varepsilon_1,\varepsilon_2}(x) = 1 - (1-x)^{\varepsilon_1 + (\varepsilon_2 - \varepsilon_1)x}$.

No closed form solution of either of the foregoing regression problems is known, not even in the simplified case of an equality parameter rather than an equality function which means that the regression variables are set equal to $\varepsilon_1 = \varepsilon_2$ and vary in common through the candidate set.

3.2 Linear equality functions of nations

Data compiled by the world bank [W] were used to compute best fit linear equality functions according to the foregoing regression scheme for the continuous and the implicit Pareto distribution. The settings $Z' = \{.1, .2, .4, .6, .8, .9\}$ and $x_i = \frac{i}{10}$ for $i = 0, \ldots, 10 = n$ were used. The regression parameters ε_1 and ε_2 varied independently from 0 to 1 with a step size of .01. The regression data are given in the following 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
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
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
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
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
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
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
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
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
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
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
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
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
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
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
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
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
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
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
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
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
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
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
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
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
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
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
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
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
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
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

The equality parameters according to the ordinary Pareto distribution for theses data have been computed earlier, see [KPR]. Their listing is included for comparability to linear equality functions. In all cases, the best fit equality functions turn out to be increasing. This means that the poor are not only absolutely poorer than the rich – which is a tautology – but also relatively; a lower income corresponds to a lower fraction of the average of all incomes above.

The equality parameters are approximately equal or lie above the mean of the lower and the upper equality parameters of the implicit Pareto distribution. The average of upper and lower equality parameter of the continuous Pareto distribution consistently underestimates the equality parameter. This applies to all nations investiaged. All parameter values are listed in the following table.

Nation	Pareto distribution	continuous Pareto distribution		implicit Pareto distribution	
	ε	ε_1	ε_2	ε_1	ε_2
Austria	0.65	0.45	0.74	0.47	0.88
Brazil	0.27	0.03	0.35	0.08	0.47
Canada	0.55	0.29	0.66	0.33	0.81
China	0.45	0.18	0.54	0.22	0.69
Czech Rep.	0.62	0.47	0.68	0.49	0.77
Denmark	0.63	0.41	0.73	0.45	0.86
Finland	0.61	0.43	0.69	0.46	0.80
France	0.54	0.28	0.64	0.33	0.78
Germany	0.59	0.37	0.68	0.41	0.80
Gr. Britain	0.50	0.24	0.60	0.29	0.74
Greece	0.54	0.30	0.63	0.34	0.77
Hungary	0.52	0.36	0.63	0.39	0.74
India	0.47	0.38	0.50	0.41	0.53
Italy	0.59	0.36	0.70	0.40	0.84
Japan	0.62	0.46	0.65	0.49	0.78
Korean Rep.	0.55	0.30	0.65	0.34	0.80
Mexico	0.33	0.10	0.40	0.15	0.50
Netherlands	0.54	0.29	0.64	0.34	0.78
Nigeria	0.35	0.15	0.42	0.20	0.51
Norway	0.62	0.43	0.70	0.46	0.82
Poland	0.53	0.33	0.61	0.35	0.74
Portugal	0.50	0.29	0.58	0.33	0.65
Russian Fed.	0.37	0.16	0.44	0.20	0.55
S. Africa	0.28	0.01	0.36	0.06	0.49
Slovakia	0.69	0.53	0.77	0.56	0.86
Spain	0.54	0.29	0.64	0.34	0.77
Sweden	0.63	0.40	0.73	0.43	0.88
Switzerland	0.53	0.28	0.64	0.32	0.79
USA	0.47	0.14	0.59	0.19	0.78
Venezuela	0.38	0.10	0.47	0.15	0.61

The interpretation of these numbers is illustrated for the implicit Pareto distribution in case of Germany. The income of the poorest approximately equals $0.41 \times 100\% = 41\%$ of the overall average income. Strictly, this is an infeasible statement since the empricial data which were used for regression do not cover "the" lowest and "the" highest income in the population. The income at the 30% quantile approximately equals $(0.41 + (0.80 - 0.41)\,0.3) \times 100\% = 52.7\%$ of the average of the 70% largest incomes of the nation.

A mere equality parameter above 0.5 might suggest non-existence of poverty according to the relative poverty notion of the Eropean Union. The shortfall of an equality function below the value of 0.5 at the lower end of the income scale and the slope of the equality function indicate the extend of relative poverty. Based on the regression data, only one nation (Slovakia) – not even one of the Scandinavian nations – appears to be free of relative poverty. This issue is made precise by relative poverty lines, see section 3.3.

The nation with largest difference between upper and lower equality parameter (slope of the equality function) both for the continuous and for the implicit Pareto distribution is USA. The nation with smallest difference between upper and lower equality parameter both for the continuous and for the implicit Pareto distribution is India. It is remarkable that these two nations exhibit the extremes since their ordinary equality parameters are equal.

The equality function of the implicit Pareto distribution lies above the equality function of the continuous Pareto distribution for all nations. This is in line with the dominance relation between the two distributions, comp. section 2.3.2. Not only values but also the slopes of the equality functions are larger under implicit Pareto distributions than under continuous Pareto distributions.

The regression error of best fits decreases drastically when replacing the one-parametric Pareto-type curves with the two-parametric continuous and implicit Pareto-type Lorenz curves. This is illustrated by the subsequent table.

Nation	Regression error		
	Pareto distribution	continuous Pareto distribution	implicit Pareto distribution
Austria	.00378	.0000428	.000130
Brazil	.00902	.000105	.000015
Canada	.00677	.0000794	.000236
China	.00842	.0000419	.000049
Czech Rep.	.00206	.000121	.000257
Denmark	.00423	.000175	.000365
Finland	.00301	.0000463	.000125
France	.00643	.0000743	.000282
Germany	.00441	.0000311	.000147
Gr. Britain	.00693	.0000253	.000164
Greece	.00586	.0000356	.000192
Hungary	.00371	$9.76 \cdot 10^{-6}$	.0000573
India	.00897	.000241	.000356
Italy	.00516	.0000854	.000272
Japan	.00209	.0000177	.0000734
Korean Rep.	.00627	.0000797	.000281
Mexico	.00668	.0000178	.000092
Netherlands	.00618	.0000847	.000313
Nigeria	.00496	.0000313	.000159
Norway	.00328	.0000604	.000164
Poland	.00449	.0000360	.000438
Portugal	.00439	.0000262	.000134
Russian Fed.	.00584	.0000370	.000219
S. Africa	.01070	.000913	.000309
Slovakia	.00206	.0000635	.000144
Spain	.00584	.0000827	.000261
Sweden	.00481	.0000810	.000198
Switzerland	.00693	.0000115	.000375
USA	.01197	.0000504	.000230
Venezuela	.00942	.0000156	.000191

The regression error of the continuous Pareto distribution is consistently smaller than that of the implicit Pareto distribution with the exceptions of Brazil and South Africa. Whenever the regression error is smaller, it is so by about one order of magnitude. The regression error of the two-parametric continuous Pareto distribution is also smaller than those of several one-parametric distributions that were obtained from modifying the equality parameter by an overconstraining method, see [KR]. The regression error of the implicit Pareto distribution is consistently smaller than that of the ordinary Pareto distribution and, again, it is so by about one order of magnitude.

3.3 Relative poverty lines

The fraction of a population whose income falls short of 50% of the overall average income can be considered as the relative poverty line according to the EU poverty definition. The relative poverty line is given as the solution x_{50} of the equation $y'(x_{50}) = 0.5$. For the continuous and implicit Pareto distributions the relative poverty lines can be computed by the respective formulas

$$0.5 = y'_{cP,\varepsilon_1,\varepsilon_2}(x_{50})$$

$$= (1 - x_{50})^{\varepsilon_1 + (\varepsilon_2 - \varepsilon_1) x_{50} - 1} \cdot \Big((1 - x_{50}) \ln(1 - x_{50}) \cdot (\varepsilon_1 - \varepsilon_2) + (\varepsilon_1 + (\varepsilon_2 - \varepsilon_1) x_{50}) \Big)$$

and

$$0.5 = y'_{iP,\varepsilon_1,\varepsilon_2}(x_{50}) = (\varepsilon_1 + (\varepsilon_2 - \varepsilon_1) x_{50}) \frac{1 - y_{iP,\varepsilon_1,\varepsilon_2}(x_{50})}{1 - x_{50}}.$$

The relative poverty line of a continuous Pareto distribution can be approximated by successive division of the function domain and selecting that half of the present interval where the derivative $y'_{cP,\varepsilon_1,\varepsilon_2}(x)$ attains the value 0.5 ("regula falsi"). The relative poverty line of an implicit Pareto distribution can be approximated from the support vertex x_i at which the left and right slopes sandwich the value 0.5. This support vertex is given by

$$y'_{iP,\varepsilon_1,\varepsilon_2}(x_i-) < 0.5 < y'_{iP,\varepsilon_1,\varepsilon_2}(x_i+).$$

The approximation then is

$$x_{50} = x_{i-1} + \frac{0.5 - y'_{iP,\varepsilon_1,\varepsilon_2}(x_i-)}{y'_{iP,\varepsilon_1,\varepsilon_2}(x_i+) - y'_{iP,\varepsilon_1,\varepsilon_2}(x_i-)} \cdot (x_{i+1} - x_{i-1}).$$

Alternatively, as done here, approximations can be obtained by sampling the domain at constant step size and selecting the argument for which the Lorenz curve derivatives come closest to the target value 0.5. A step size of 0.001 was chosen for the continuous Pareto distribution and a step size of 0.01 was chosen for the implicit Pareto distribution. In all computations the lower and upper equality parameters are chosen to be the best fit values obtained in section 3.2. The resulting relative poverty lines are given in the subsequent table.

Nation	Relative poverty line	
	continuous Pareto distribution	implicit Pareto distribution
Slovakia	0.000	0.00
Czech Rep.	0.043	0.02
Austria	0.058	0.04
Japan	0.061	0.02
Norway	0.084	0.06
Finland	0.086	0.07
Denmark	0.095	0.07
Sweden	0.103	0.10
Germany	0.137	0.13
Italy	0.139	0.14
Hungary	0.161	0.17
Poland	0.188	0.21
Korean Rep.	0.189	0.21
Canada	0.190	0.21
Greece	0.196	0.21
Netherlands	0.197	0.21
Spain	0.197	0.21
Portugal	0.197	0.26
France	0.202	0.22
Switzerland	0.202	0.23
India	0.213	0.21
Gr. Britain	0.235	0.27
USA	0.271	0.32
China	0.283	0.34
Venezuela	0.339	0.41
Russian Fed.	0.347	0.41
Nigeria	0.363	0.43
Mexico	0.385	0.46
S. Africa	0.410	0.51
Brazil	0.420	0.51

Values almost coincide in some cases. When they differ, no approximation of the relative poverty line can be guaranteed to be closer to the true value than the other approximation and no systematic over- or underestimation is known. Value discrepancies result from the derivatives of the Lorenz curves being flat towards the lower end of the curves.

The table is sorted by increasing poverty lines of the continuous Pareto distribution. Ties such as between France and Switzerland are broken in favour of increasing poverty lines of other distribution. Remaining ties such as between the Netherlands and Spain are broken arbitrarily.

The relative poverty lines of the implicit Pareto distribution tend to vary more than those of the continuous Pareto distribution. At the low end of the poverty scale the implicit Pareto distribution has smaller poverty lines than the continuous Pareto distribution which is illustrated by the first nine entries of the foregoing table. When poverty increases, the relation reverses with the exception of India.

3.4 World income distribution

The individual income distributions of various nations can be aggregated to the world income distribtion when "weighted" by national gross incomes and population sizes. This results in the following regression support for more than 150 nations of the world, see [StK].

"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

Regressions as for individual nations result in the following best fit values.

"Nation"	Pareto distribution	continuous Pareto distribution		implicit Pareto distribution	
	ε	ε_1	ε_2	ε_1	ε_2
World	0.12	0.00	0.15	0.00	0.23

The lower equality parameter of 0.00 for the continuos and implicit Pareto distributions must not be considered as zero, but, within the numerical precision of 0.01, the true values are closer to 0 than to 0.01. The regression errors of the three Pareto distributions fitted to the world data are as follows.

"Nation"	Regression error		
	Pareto distribution	continuous Pareto distribution	implicit Pareto distribution
World	.0128	.0055	.0037

The smallest regression error is attained by the implicit Pareto distribution in analogy to few cases of the nation analysis, namely Brazil and South Africa.

4 Conclusion

Linear equality functions have been developed and fitted to real data. All in all, the continuous Pareto distributions exhibits a better fit than the implicit Pareto distribution. Relative poverty lines are derived for both distributions with no consistent over- or underestimation of one distribution by the other. Future investigations will have to address non-linear equality functions.

References

[BouMe] 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, also under www2.soc.hawaii.edu/econ/workingpapers/992.pdf.

[ChoGri] Chotikapanich, D., Griffiths, W.E., "Estimating Lorenz curves using a Dirichlet Distribution", Working paper 110, University of New England, Armidale, 1999.

[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.

[EmKlMi] Embrechts, P., Klüppelberg, C., Mikosch, T., "Modelling extremal events", Springer, Berlin, 1997.

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

[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.

[K] Kämpke, T., "Overconstrained ordinary differential equations", International Mathematical Journal 2, 2002, p. 1125-1139.

[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.

[KR] Kämpke, T., Radermacher, F.J., "equality analysis by functional approach", 2002, submitted.

[Lit] Litchfield, J.A., "Inequality: methods and tools", Discussion paper, Worldbank, Washington, 1999, www.worldbank.org/poverty/inequal/methods/litchfie.pdf.

[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., Slottje, 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.

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

[W] Worldbank, "World Development Report 2000/2001", Washington, 2001, www.worldbank.org/poverty/wdrpoverty/report/index.htm.