The use of mean values vs. medians in inequality analysis

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

Abstract

Poverty lines that are proportional to either the mean value or the median of income distributions are compared by statistical properties and in the light of poverty axioms. Poverty lines are extended from all incomes such that any particular income is considered as smallest of all larger incomes. This induces classes of income distributions with the distributions for the median being much more unequal than for the mean value.

Key words: income distribution, Lorenz curve, poverty line.

1 Introduction

The central question of this investigation is whether a poverty line should be proportionate to the mean value or to the median of an income distribution. While for most people the natural reference is the mean value, official standards are developing more and more in the direction of using the median. It is argued that this approach is of political significance. It allows to camouflage asymmetries that when fully transparent, might have political consequences.

The dispute whether the mean value or the median is an appropriate indicator of the average of an income distribution has a long history. Also, it may seem that this dispute is a technicality that keeps statisticians, econometricians and others busy without leading anywhere. But favouring one measure over the other can matter; this so in particular, when average income is used to measure poverty.

As a refresher, the mean value of any data set is computed by adding all data and dividing by the number of the data. The median is defined as the middle value which means that 50% of the data are less or equal and 50% of the data are larger or equal. The data set $\{1000, 1700, 3000, 4000, 10300\}$ has the mean value $\frac{1000+1700+3000+4000+10300}{5} = 4000$ and the median 3000. If the set has an even number of data, the median is the mean value of the two middle values: the data set $\{1000, 1700, 3000, 4000, 10300, 20000\}$ has the median $\frac{3000+4000}{2} = 3500$.

The European Union relates its definition of individual poverty to the median. Accordingly, a person is considered poor if his income falls short of 60% of the median of all incomes. On the other hand, the so-called objective 1 areas in the terminology of the European Regional Development Fund ERDF were defined by mean values. A region was considered poor and qualified for funding, if its per capita GDP (gross domestic product) fell short of 75% of the EU mean value.

The Organization for Economic Co-operation and Development OECD often uses the EU poverty line of 60% of the median while some others, like United Nations organiztions, occasionally use 50% of the median for poverty assessment. The US Census Bureau uses the median as a plain indicator of the average income [USC].

Strictly, income and poverty are related to households instead of individual persons in order to avoid accounting many persons to have zero income when they actually benefit from income of other household members. This is typically achieved by so-called equivalence numbers. Household corrected incomes refer to the mean value or the median in the same way as uncorrected incomes so that household correction is ignored in the sequel.

The German federal government in its first "report on poverty and wealth" used the four combinations of mean value and median as well as of the 50% and the 60% fraction. Mean values were larger than medians but 60% of the median was still larger than 50% of the mean value over the complete reporting period (1973-1998), [Bu1, p. 25]. A similar finding is reported for most EU nations [No, p. 6]. In the meantime, the "report on poverty and wealth" is harmonized with the EU standard so that the 60% fraction of the median is used exclusively [Bu2].

Besides the mean value of all incomes typically being larger than the median of all incomes, a closer view may show that mean incomes and median incomes are very similar in lower percentiles but mean values clearly exceed medians in upper percentiles of an income distribution. This was found to be true for US income distributions [Rey].

The decision in favour of mean or median should not adhere to pushing up or down the poverty rate in official statistics possibly following opportunistic motivations. In particular, the decision should not be based on adjusting the fraction to 50%, 60% or else. Rather, sound arguments should lead to the decision for either mean value or median and, though more arbitrary, to the fraction by which it is used for a poverty line.

A fairly common argument in favour of the median is that the average of an income distribution should account for its asymmetry. An income distribution tends to have more low than high incomes so that it is skewed towards the upper part. Income distributions have so-called thin tails at their high ends. While the median is unaffected by few large incomes, the mean value is not. This robustness, which is explained precisely below, is claimed to be the salient property of the median warranting it to represent the statistical average.

When it comes to finding a poverty line, another argument in favour of the median is occasionally used which is less clear and almost tautological. The argument is that "... mean income produces excessively high poverty rates" [SaSm]. But rating a poverty notion as 'excessive' is only meaningful if poverty can be made operational otherwise. In principle, this can be achieved by baskets consisting mainly of food items. However, the different strength of association between affordability of such baskets and fractions of either median or mean income seems to be difficult to capture and rarely approached. On the other hand, low absolute poverty lines like $1, $2 and even $3 per day seem to be widely accepted for "poor" countries as well as a threshold like $10787 per annum for a one person household in the US.

Poverty measures, in contrast to poverty lines, do not (only) intend to count the poor but to denote the distribution of their incomes. Poverty measures typically require a poverty line to be specified which the poverty measures themselves do not do. The present work is complementary to poverty measures in that the most frequent candidates, namely mean value and median-based poverty lines are analyzed.

The dispute between mean value and median is reasonably approached by comparisons – not of numerical values but of effects from distributional changes. For example, what happens if a certain section of all incomes, say the top 20%, rise and all others nominally remain unchanged? This has been a stylized pattern in the globalized economy for the last 30 years. The median does not change, the mean value increases and, if this trend is going to continue for some time, the portfolio of products and services can reasonably be speculated to shift towards the high income group. Put differently, the portfolio of products and services available to the "below average income person" shrinks, at least in relative terms. This means nothing else than the "below average persons" become relatively poorer or – as often called – deprived. The mean value would account therefore, because low incomes, though assumed to remain nominally equal, become relatively smaller when compared to the new, larger mean value. The median would hide this widening gap.

In the following, income distributions will be described, as often, by Lorenz curves and by their derivatives which are denoted as income curves or Lorenz-income curves. Lorenz curves and income curves are defined over the interval from zero to one. The income curve accounts for increasingly sorted incomes. At any point x, the income curve denotes the relative income level such that $100 \cdot x\%$ of all incomes are smaller and $100 \cdot (1-x)\%$ of all incomes are larger. The total income of a population is thereby normalized to one. An income curve is always increasing.

Income curves should not be confused with density functions over absolute incomes. These describe absolute or relative frequencies of absolute income levels.

The Lorenz curve at any point x denotes, as usual, the relative cumulative income of the $100 \cdot x\%$ lowest incomes. The cumulative income being relative means that it is devided by the sum of all incomes. To account for those without own income but living in households with income, actual income values are typically adjusted to household sizes by equivalence numbers. A Lorenz curve F is always increasing, convex and normalized to the values $F(0) = 0$ and $F(1) = 1$.

The remainder of this paper is organized as folows. Estimation properties of the mean value and the median will be discussed in section 2. Particular properties in the context of income distributions are summarized in section 3 in the light of poverty axioms. The use of mean values and medians is extended from all incomes to segments of incomes greater than certain bounds in section 4. This analysis reveals that the extension of median-based poverty lines leads to drastically larger inequality than the extension of mean value-based poverty lines. Section 5 draws some conclusions.

2 Statistical properties

The mean value is the least square approximation of all income data while the median is the least absolute approximation. This means that for any data set $\{x_1, \ldots, x_n\}$

$$argmin_x \sum_{i=1}^{n} (x_i - x)^2 = \mu$$

with $\mu = \frac{x_1 + \ldots + x_n}{n}$ and

$$argmin_x \sum_{i=1}^{n} |x_i - x| = med$$

with *med* denoting the median of the data set. It should be clear that the mean is more sensitive than the median to changes of large data. Even more, arbitrary large changes do not affect the median as long as no more than 50% of the data change. This is known as breakdown point.

The breakdown point of an estimator is the proportion of data which may be driven towards infinity without the estimator itself being driven beyond any limit [LoRo]. A mean value can be driven beyond any limit by just one datum itself tending towards infinity and all other data remaining unchanged. So, the mean value has the smallest possible breakdown point zero.

In contrast, the median is considered robust since it has a positive breakdown point. The breakdown point of the median equals $\lfloor \frac{n-1}{2} \rfloor / n$; $\lfloor x \rfloor$ is the so-called floor of x which is the greatest integer less or equal to x. For the even number $n = 8$ the breakdown point is $\lfloor \frac{7}{2} \rfloor / 8 = 3/8 = 0.375$ and for the odd number $n = 7$ the breakdown point is $\lfloor \frac{6}{2} \rfloor / 7 = 3/7 = 0.4286$; in both cases three data may tend towards infinity without the median tending towards infinity.

The breakdown points approach 0.5 for ever growing data sets which means that almost 50% of the data can increase beyond any limit without driving the median beyond any limit. The median is not even slightly changed by such growth if those values grow which were the largest in the first place.

Robustness does not imply that the median always changes less than the mean value under equal alterations of incomes, not even if less than 50% of all incomes are altered. Such an "instability" effect occurs when few incomes above the current median sharply drop below that median and all the other incomes below lie so far below that they cannot support the median decline.

The robustness property may be unexpected when considering the different features of mean and median for the first time. Robustness of the median cannot be disputed so that the median better than the mean value copes with errors, in particular with uneven errors. Uneven errors comprise the hiding of incomes at the high end and omissions at the low end of the income scale. But, considering complete and honest input to statistics, this effect does not settle the issue whether robustness is sufficient reason to select the median as indicator of average income.

3 Analysis from plain statistics

The median-based poverty notion of the EU can be altered so as to refer to the mean value. Obviously, this means that a person would be considered poor if his income falls short of 60% of the mean value of all incomes. Plain statistics provide no ground to decide whether the mean value or the median is more appropriate for measuring poverty. Both may lead to unexpected non-monotonicity effects.

3.1 Median

While the mean value is independent of any redistribution of income – be it tax induced or else – the median is not. Redistribution from the middle to the top may even pretend a decrease in poverty. This strange non-monotonicity effect is possible if middle incomes are considered to begin right above the median. The effect is reported to actually have occured for income data of New Zealand in the 1990's [Ea].

For the sake of illustration, the set of five data from the introduction is modified to $\{1000, 1700, 2500\,(= 3000 - 500), 3500\,(= 4000 - 500), 11300\,(= 10300 + 1000)\}$. The mean value remains unchanged at 4000, but the median decreases from 3000 to 2500. When using the median-based poverty definition of the EU, the poverty line decreases from $0.6 \cdot 3000 = 1800$ to $0.6 \cdot 2500 = 1500$. Thus, the number of poor decreases from two to one simply by income redistribution from the middle to the top.

The poverty line in terms of 60% of the mean value remains at $0.6 \cdot 4000 = 2400$. This would leave the number of poor at two before and after redistribution.

The stated redistribution increases the original income disitribution in the sense of the danger order [MSt, p. 23]. An income distribution X is understood to be less dangerous than an income distribution Y if

1. $EX \leq EY$ (the order applies to the mean values) and
2. there is a point x_0 such that $F_X(x) \leq F_Y(x)$ for all $x < x_0$ and $F_X(x) \geq F_Y(x)$ for all $x \geq x_0$. (The distribution function of an income distribution X with n incomes is understood as usual by

$$F_X(x) = \frac{\text{number of incomes } \leq x}{n}.)$$

Large "risks" are less probable under X than under Y which explains the terminology.

The danger order is also known as Karlin Novikoff criterion that is sufficient for the increasing convex order of two distributions. The observed non-monotonicity for a situation with $n = 5$ incomes means that an income distribution with more variability in terms of "danger" as well as "increasing convexity" may have the smaller median at equal mean values.

3.2 Mean value

A similar non-monotonicity effect is exhibited by the mean value in case some of the incomes below the median rise and all other incomes remain unchanged such that the median remains unchanged. Then the mean value rises and the number of poor may rise as well.

This effect is illustrated by an example which is only slightly more complicated than the foregoing. The data set is initially assumed to consist of the incomes $1300, 1500, 2300, 3000, 4000, 5000, 9500$. The mean value is 3800 so that the 60% poverty line lies at $0.6 \cdot 3800 = 2280$. This means that two individuals are poor.

Now only the two poor receive an increase so that the median of 3000 remains unchanged. The new income values are $1500\,(= 1300 + 200), 2000\,(= 1500 + 500), 2300, 3000, 4000, 5000, 9500$. The new mean value is 3900 and its 60% poverty line is at $0.6 \cdot 3900 = 2340$. This means that now three individuals are poor.

3.3 Polarization

When a given income distribution becomes more polarized, both mean values and medians tend to change. For example, when the lower two thirds of all incomes decrease by 5% and the upper third increase by

5%, then the median also decreases by 5%. Even more, scaling within the lower two thirds leaves the number of all incomes below any fixed fraction of the median unchanged.

But the mean value decreases by less than 5% or even increases depending on the specific data. Thus, more incomes tend to fall below any fixed fraction of the mean value which invariably increases the poverty headcount.

Polarization was used so far in an informal sense only. Well established rigorous notions of polarization for income distributions are not applied here since they critically rely on scaling parameters as does the Esteban Ray index [ER]. Varying the scaling parameter can lead to anything between unimodularity (one local maximum) indicating no polarization and an abundance of peaks (multimodularity) indicating pronounced polarization. Neither, the Wolfson bipolarization index which involves both the median and the mean value is not used here since it always refers to a fixed bisection into the lower and upper 50% income segments. These can be generalized to two other segment sizes [Rod].

Polarization is generally understood to be a measure of income clusters and, so, is different from inequality in that a certain inequality may go along with large polarization and large inequality may have little polarization as indicated in figure 1. But polarization always entails inequality.

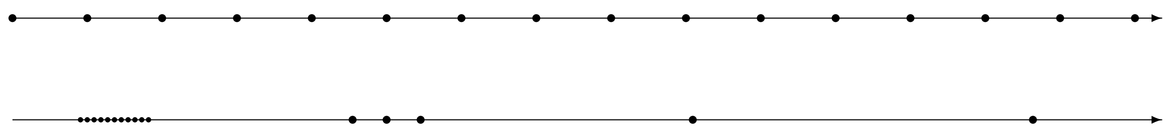

Figure 1: Two hypothetical sets of $n = 16$ income data. The upper distribution has a certain inequality and adjacent incomes are equidistant so that the distribution has no apparent polarization. The lower distribution has less inequality (masured by the Gini index or by the difference between maximum and minimum incomes) but clearly more polarization.

One approach to measuring polarization of an income distribution is by formal classification methods. Without any parameter specification, the incomes can be split into three clusters whose centers then serve for further analysis. For example, the median falling into the same or another cluster than the mean values is an indicator of polarization.

Clustering the sorted incomes $x_1 \leq \ldots \leq x_n$ into, say, a lower, middle and upper segment I_l, I_m and I_u follows the minimization of squared intra-class distances in a straightforward manner. Distances between members of the same class are made small without explicitly driving up distances between members of different classes.

$$\min_{I_l, I_m, I_u \text{ partition of } \{1,\ldots,n\} \text{ and } a_l, a_m, a_u \in \mathbb{R}} \sum_{i \in I_l} (x_i - a_l)^2 + \sum_{i \in I_m} (x_i - a_m)^2 + \sum_{i \in I_u} (x_i - a_u)^2.$$

Given any partition, the best class centers are computable as the cluster centroids

$$a_l = \frac{1}{|I_l|} \sum_{i \in I_l} x_i, \; a_m = \frac{1}{|I_m|} \sum_{i \in I_m} x_i, \; a_u = \frac{1}{|I_u|} \sum_{i \in I_u} x_i.$$

As a consequence from these computations, the clustering problem reduces to the discrete problem of finding an optimal triplet of sets. For sorted incomes it suffices to consider clusters of consecutive indices $I_l = \{1, \ldots, l\}$, $I_m = \{l+1, \ldots, m\}$ and $I_u = \{m+1, \ldots, n\}$. The centroids of all such clusters are computed for all $O(n^2)$ triplets and the triplet with minimum sum of intra-class distances is retained as best clustering of a set of income data.

The mean value of the income distribution is related to the cluster centroids by the weighted average

$$\mu = \frac{|I_l|}{n} a_l + \frac{|I_m|}{n} a_m + \frac{|I_u|}{n} a_u.$$

Any value like the mean value, the median or else is naturally assigned to that segment whose centroid is closest. So, the median lies in the lower segment if $med \leq \frac{a_l+a_m}{2}$ and it lies in the middle segment if $\frac{a_l+a_m}{2} \leq med \leq \frac{a_m+a_u}{2}$ where ties for equality are broken arbitrarilty. The mean value and the median are actually assigned to the middle segment in all foregoing numerical examples but things are different in the subsequent slightly more complex example.

The $n = 16$ incomes $x_1 = 1000, x_2 = 1100, x_3 = 1200, \ldots, x_{11} = 2000, x_{12} = 5000, x_{13} = 5500, x_{14} = 6000, x_{15} = 10000, x_{16} = 15000$, see lower section of figure 1, result in the three clusters $I_l = \{1, \ldots, 11\}$, $I_m = \{12, 13, 14\}$ and $I_u = \{15, 16\}$ with centroids

$$a_l = 1500,\ a_m = 5500, a_u = 12500.$$

The mean value 3625 is assigned to the middle section since $\frac{a_l+a_m}{2} = 3500 < 3625 < 9000 = \frac{a_m+a_u}{2}$ but the median 1750 is assigned to the lower section since $1750 < 3500 = \frac{a_l+a_m}{2}$.

As an alternative, measuring the polarization of income distributions can be approached by schemes for function approximating with piecewise constant functions, called quantization [K], applied to Lorenz curves, their derivatives or to frequency functions of absolute incomes. Classes are thus formed by the segments over which a best approximating function is constant. The number of segments is thereby set externally. In analogy to clustering, the number can be set to three. Mean values and medians then lying in the same or in different income segments is a polarization indicator.

3.4 Context

It were of some help if either the mean or the median income showed a stronger association with other statistics, for example with health statistics. But such an association seems to be weak overall [NaDo] denying any decision in favour of one of the two averages. This is nonwithstanding the circumstance that life expectancy is generally increasing in per capita GDP across nations [Ma, p. 15].

3.5 Poverty axioms

Most axioms for poverty measures are not satisfied by the poverty lines of either median or mean value. The reason is that these axioms aim at poverty measures that do not only separate poor from non-poor but also quantify poverty [BeLi]. In particular, the three axioms by Sen are satisfied by mean value or median-based poverty lines, if at all, to a marginal extend. More precisely, due to quantization effects, the axioms may not be violated for small or few changes but they will be for large or many changes stipulated in the axioms.

The focus axiom states that the poverty measure remains unchanged when a non-poor receives an increase in income. The increase is meant to be external (no redistribution). The mean value then changes but the median does not when actually only a single non-poor receives the increase. The focus axiom is then satisfied for then median.

The monotonicity axiom requires that an increase in income of any single poor leads to a decrease of the poverty measure. In fact, this is the strong monotonicity axiom, where the weak monotonicity axiom requires that an increase in income of a poor leads to a decrease of the poverty measure if the poor individual remains poor after the increase. Mean value-based poverty lines satisfy this for suitable values but violate it for others as exemplified above. Median-based poverty lines do not violate the axiom for suitable values.

The transfer axiom refers to transfer among the poor only, namely from a poor to a less poor. This counter-intuitive transfer should increase the poverty measure. The mean value and the median do not satisfy this axiom.

Other than Sen's axioms include that of scale invariance. This axiom states that a poverty index remains unchanged if the incomes of the poor and the poverty line are scaled by the same factor. This axiom need not be satisfied by mean and median-based poverty lines, because possible changes of non-poor incomes are left untouched. But the scale invariance axiom is satisfied by both mean and median if modified so that all incomes are scaled by the same factor.

It is obvious that the foregoing poverty axioms do not segregate the median from the mean value. This applies to further axioms as well without being elaborated here.

3.6 Further axiomatic features

The intuition of a possible axiom for poverty lines should be that a rising gap between those who are informally understood as rich and poor increases the poverty line. Widening inequality is thereby understood as Lorenz-dominance as shown in figure 2.

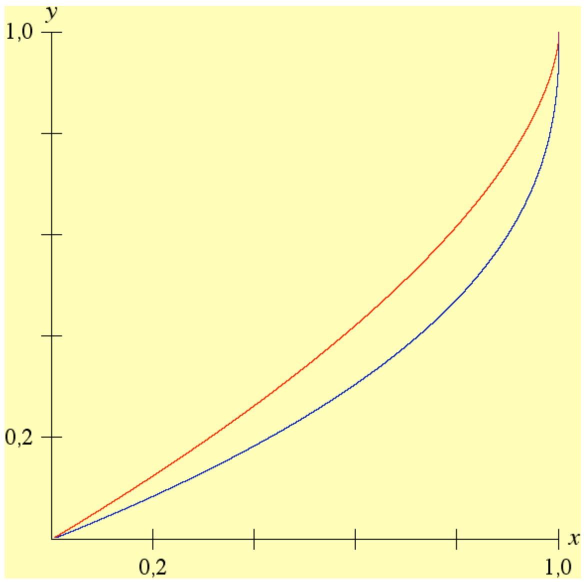

Figure 2: When all points of one Lorenz curve lie below (or at equal level of) another Lorenz curve, the lower curve is understood to dominate or "Lorenz-dominate" the upper; it exhibits more inequality than the upper curve.

A possible dominance axiom would then be as follows. When two income distributions have the same mean value such that one Lorenz-dominates the other, then the number of poor according to the Lorenz-dominating distribution should be increasing (non-decreasing) as compared to the other distribution.

The requirement that the mean values be identical cannot be dropped without any replacement. A counter example to mean value-based poverty lines satisfying a hypothetical dominance axiom that stipulates Lorenz-dominance only is that of section 3.2.

But even if the mean values are identical, the possible dominance axiom in its present form may be violated by mean value-based poverty lines. The same applies to median-based poverty lines which is illustrated by the example from section 3.1.

Hard upper bounds on the size of the poverty segment are avoided by merely stipulating it in an "axiom" as follows.

Unbounded poverty axiom. Any poverty line should be defined irrespective of the underlying income distribution allowing that the proportion of the poor is a-priori unbounded.

This axiom is satisfied by mean value-based poverty lines but not by median-based poverty lines which conceptually restrict the proportion of the poor to 50%. The latter can only be circumvented by allowing a fraction of more than 100% for setting the poverty line. Considering the typical 50 or 60 per cent fractions, this seems unrealistic.

A restriction of the poor segment to half of the population is unacceptable for distributions that are so uneven as those of some regions in the world and the world as a whole (each nation weighted by its population rather than each nation counting "one"). For example, the finding "... on the \$3/day standard, 64% of Africans were poor in 1998" [Le] is impossible in terms of the median. The median of this income distribution obviously is lower than \$3 per day so that a portion of the 64% lowest income segment, technically, were not poor.

Accepting a median-based poverty line is reasonable, at best, for somewhat balanced income distributions as they are typical for developed nations. But this would lead to a double standard. The reason is that median-based poverty lines do not make sense for quite unbalanced income distributions. Such distributions are not fabricated for the sake of an argument; they exist.

4 Extending mean and median-based poverty notion

A spreading technique will now be introduced for not only relating the smallest income to all others but each income to all larger ones. This technique will demonstrate the difference between mean value and median when extended to the whole income range.

The idea is that of self-similarity. The whole income distribution and all its truncated income distributions are considered in exactly the same way, where truncation always refers to incomes below any threshold. So, emphasize is given to the view that poverty is not absolute but strictly relative and, less widespread, that inequality may be more or less homogenous throughout the whole income distribution.

4.1 Mean value-based extension

Altering, as above, the median-based poverty notion of the EU so as to refer to the mean value allows to extend it in two ways. First, the fraction of 60% can be set to a general, yet unknown value. Second, the perspective can be generalized from the very poorest to all others by considering every individual as the poorest among all with higher incomes. An income curve for which each income is equal to the fraction ε of the mean value of all larger incomes satisfies the differential equation

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

for $0 \leq x < 1$ and so-called equity parameter $0 < \varepsilon \leq 1$. The manifold of all solutions of the linear inhomogenous differential equation admits the closed form representation

$$F(x) = F_\varepsilon(x) = 1 - (1 - x)^\varepsilon.$$

The manifold of these Lorenz curves is sketched in three dimensions in figure 3. The differential equations approach to Lorenz curves has been introduced earlier [KPR], modified to other differential equations [KR] and regressions have been computed that lead to actual values of the equity parmeter for nations and some regions of the world. For support points (x_i, y_i), $i = 1, \ldots, n$ taken from [Wo], the ordinary least squares regression problem

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

leads to the best fit equity parameters as stated in the next table.

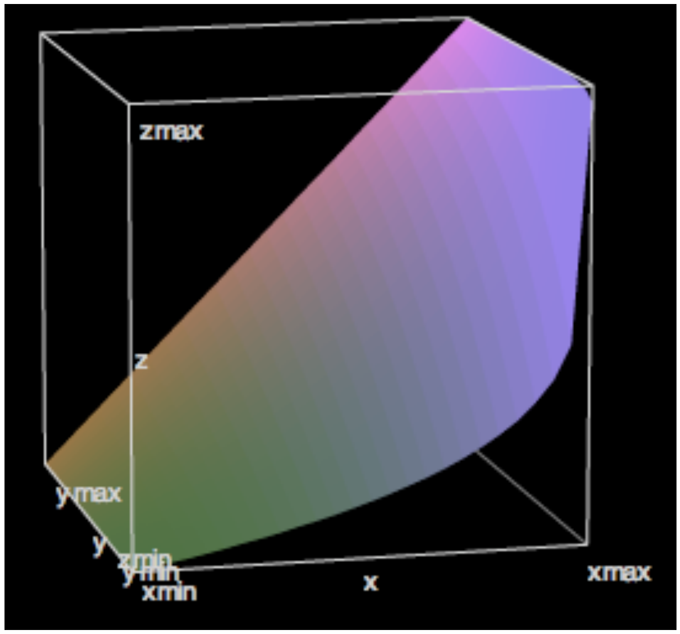

Figure 3: Manifold of Lorenz curves forming a 3D surface. The equity parameter is assigned to the axis which is almost perpendicular to the drawing plane. All curves are smooth but edges result from discretization. (Plot created by the on-line function plotter [Kas].)

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

4.2 Median-based extension

Also, the median-based poverty notion of the EU can be extended by considering other fractions than 60% and by generalizing the perspective from the very poorest to all others. Again, the latter is achieved by considering every individual as the poorest of all with higher incomes.

An income curve for which each income is equal to the fraction η of the median of all larger incomes satisfies the functional equation

$$g(x) = \eta \cdot g(\frac{x+1}{2})$$

for $0 \leq x < 1$ and parameter $0 < \eta \leq 1$. The median of all incomes over the interval $[x, 1]$ is simply chosen as the income at the midpoint $\frac{x+1}{2}$. For obvious reasons the single parameter η of the functional equation is denoted as median parameter. It is not the median itself and, neither, is proportional to the median of all incomes.

All increasing and continuous solutions of the one-parametric functional equation admit the closed form representation

$$g(x) = g_\eta(x) = c_\eta \cdot \eta^{\frac{\log(1-x)}{\log 2}}.$$

Deriving the formula for this manifold can be done by intricate inductive and continuation arguments which both are omitted here. The constant c_η must be set such that the income curve is normalized to $\int_0^1 g_\eta(x)dx = 1$ and the base of the logarithms does not matter since any fraction of logarithms is independent from their common base.

Integration leads to the Lorenz curve formula

$$G_\eta(x) = 1 - (1-x) \cdot \eta^{\frac{\log(1-x)}{\log 2}}.$$

The derivative of this curve implies the formula $c_\eta = 1 + \frac{\log \eta}{\log 2}$ for the normalization parameter. Sample values of the normalization parameter are given in the following table.

η	c_η
1.0	1.0000
0.9	0.8480
0.8	0.6781
0.7	0.4854
0.6	0.2630
0.5	0.0000
0.4	< 0

It now becomes obvious that, in contrast to the initial domain, the median parameter must not attain values between zero and 0.5. Otherwise, the normalization parameter becomes zero or negative. Negative values mean that the income curve is not integrable: the area under the curve is ∞. This results in the essential and probably surprising restriction

$$\eta > 0.5.$$

The restriction alone, which does not have a counterpart for equity parameters, indicates that the transition from the mean value to the median has quite some consequences. Interestingly, the EU poverty definition uses the median parameter $\eta = 0.6$ which satisfies this restriction though it does not extend the poverty consideration from the smallest to all incomes.

Income distributions of the type $g_\eta(x)$ will be denoted as med-curves or median-curves. The medians over all incomes according to the med-curves are the values

$$g_\eta(0.5) = c_\eta \cdot \eta^{\frac{\log(0.5)}{\log 2}} = c_\eta \cdot \frac{1}{\eta} = (1 + \frac{\log \eta}{\log 2}) \cdot \frac{1}{\eta}.$$

Sample values of the median are as follows.

η	$g_\eta(0.5)$
1.0	1.0000
0.9	0.9422
0.8	0.8476
0.7	0.6935
0.6	0.4384
0.5	0.0000
0.4	< 0

The data show that the median sharply decreases when the median parameter decreases. A median curve with parameter of 0.6 has a median income of only 43.84% of the mean value.

4.3 Comparisons by single values

A comparison of income curves and the respective Lorenz curves is given in figures 4 and 5 respectively. The curves refer to an equity parameter and a median parameter both of value 0.6 and include the equity parameter 0.5. The median-based curves exhibit a much larger inequality than the equity-based curves.

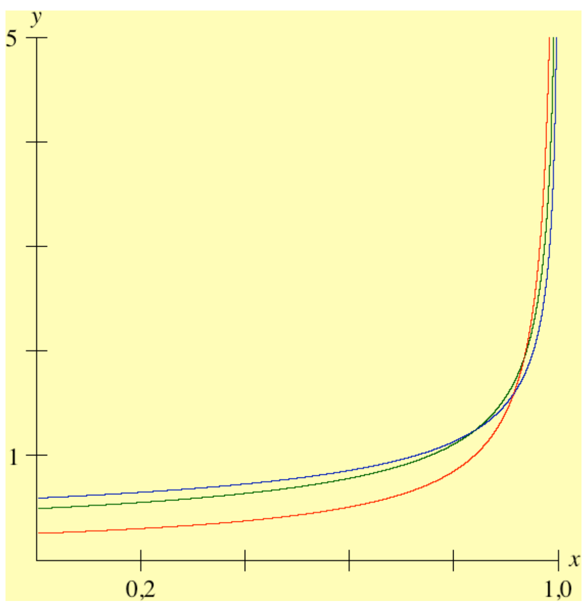

Figure 4: The median-based income curve $g_{0.6}(x) = 0.2630 \cdot 0.6^{\frac{\log(1-x)}{\log 2}}$ lies below the equity-based income curves $f_{0.6}(x) = 0.6 \cdot (1-x)^{0.6-1}$ and $f_{0.5}(x) = 0.5 \cdot (1-x)^{0.5-1}$ for over 80% of the population. The fourth possible curve, the 50% median-based curve, does not exist.

As an example, the upper permile disposes of $1 - F_{0.6}(0.999) = 0.015849$ of all incomes which is about 1.58% according to the equity-based distribution. But according to the median-based distribution this share becomes $1 - G_{0.6}(0.999) = 0.1626$ which is a stunning 16.26% of all incomes. This is more uneven than the income disttribution of the US, where a 16% share of all incomes is earned by the top 1 per cent – not 1 permile [Sh].

The numerical relation that was encountered in some empirical studies, namely that 60% of the median exceeds 50% of the mean value of all incomes can be computed to occur in equity-based distributions whenever the equity parameter exceeds 0.6570, see figure 6. Empirically, lower equity parameters also admit this phenomenon since distributions with best fit equity parameters such as those given in section 4.1, are approximations of the true income distributions only.

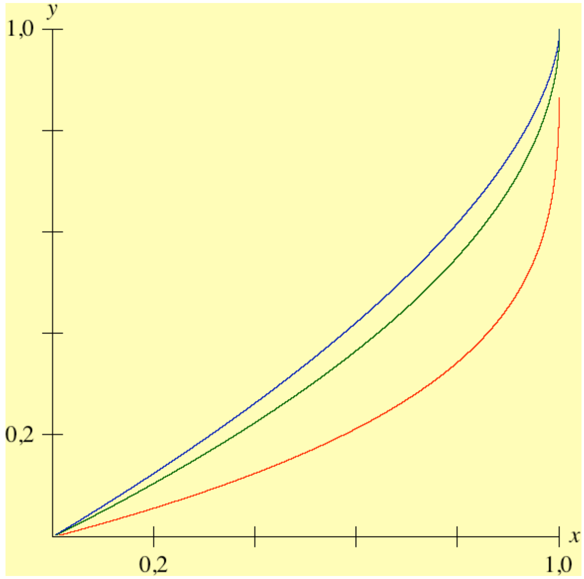

Figure 5: The median-based Lorenz curve $G_{0.6}(x) = 1 - (1 - x) \cdot 0.6^{\frac{\log(1-x)}{\log 2}}$ lies completely below the equity-based Lorenz curves $F_{0.6}(x) = 1 - (1 - x)^{0.6}$ and $F_{0.5}(x) = 1 - (1 - x)^{0.5}$. Again, the fouth possible curve, the 50% median-based Lorenz curve, does not exist.

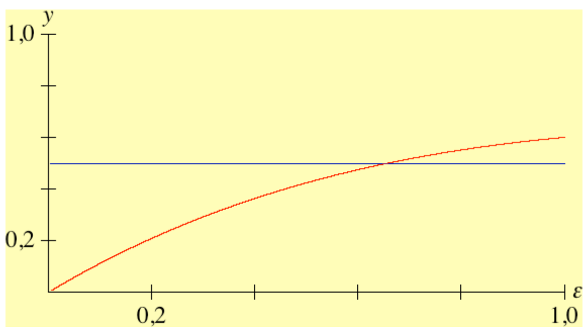

Figure 6: 60% of the median as function of the equity parameter and 50% of the mean value which is constantly 0.5. The intersection point lies at $\varepsilon = 0.6570$.

4.4 Coincidence

The med-curves and the mean value curves are identical, it only takes a parameter transformation to obtain equality. An equity-based curve with normalization parameter c_η instead of an equity parameter ε results in a median-based curve with its original median parameter.

$$
\begin{aligned}
F_{c_\eta}(x) &= 1 - (1 - x)^{c_\eta} \\
&= 1 - (1 - x)^{1+\frac{\log \eta}{\log 2}} \\
&= 1 - (1 - x) \cdot (1 - x)^{\frac{\log \eta}{\log 2}} \\
&= 1 - (1 - x) \cdot \eta^{\frac{\log(1-x)}{\log 2}} \\
&= G_\eta(x).
\end{aligned}
$$

The second last equation follows from the general identity $a^{c \cdot \log b} = b^{c \cdot \log a}$ which itself follows from taking logarithms on both sides leading to $c \cdot \log b \cdot \log a = c \cdot \log a \cdot \log b$. This implies

$$(1-x)^{\frac{\log \eta}{\log 2}} = \eta^{\frac{\log(1-x)}{\log 2}}.$$

The equality of the two Lorenz curves can also be expressed for equity parameters

$$F_\varepsilon(x) = G_{\eta(\varepsilon)}(x).$$

Equity parameters are therefore transformed to median parameters by the formula $\eta(\varepsilon) = 2^{\varepsilon-1}$. Median parameters must always be chosen larger than equity parameters to obtain the same distribution. Some corresponding values are given in the following table.

ε	$\eta(\varepsilon)$
1.0	1.0000
0.9	0.9330
0.8	0.8706
0.7	0.8123
0.6	0.7579
0.5	0.7071
0.4	0.6598
0.3	0.6156
0.2	0.5743
0.1	0.5359
– (0.0)	– (0.50)

The table is to be read as follows. When each income is equal to, say, 30% of the mean value of all larger incomes and when, according to a possibly different distribution, each income is equal to 61.56% of the median of all larger incomes, then the two distributions are actually equal as shown in figure 7.

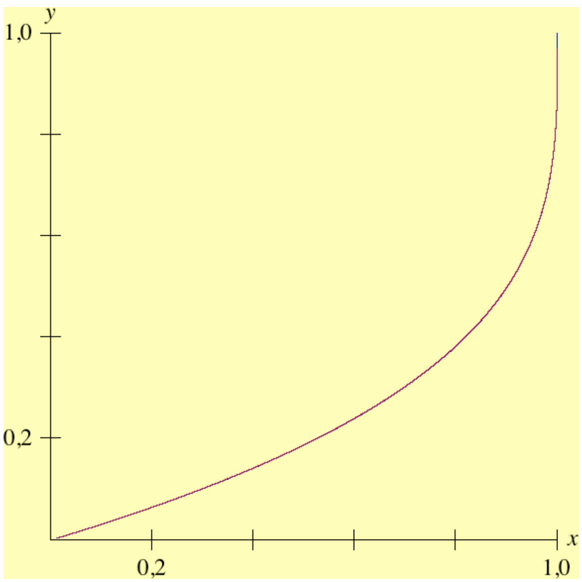

Figure 7: The Lorenz curves with equity parameter $\varepsilon = 0.3$ and median parameter $\eta = 0.6156$ are identical.

The relation between the transformations from equity parameter to median parameter and vice versa are summarized as follows.

$$
\begin{aligned}
F_\varepsilon(x) &= G_{\eta(\varepsilon)}(x) \\
F_{c_\varepsilon}(x) &= G_\eta(x)
\end{aligned}
$$

The best fit equity parameters as stated before togeher with their corresponding best fit median parameters are listed in the following table.

Nation	ε	$\eta(\varepsilon)$
Austria	0.6494	0.7843
Brazil	0.2778	0.6062
Canada	0.5525	0.7333
China	0.4464	0.6813
Czech Rep.	0.6173	0.7670
Denmark	0.6289	0.7732
Finland	0.6135	0.7650
France	0.5376	0.7258
Germany	0.5882	0.7517
Gr. Britain	0.5025	0.7083
Greece	0.5376	0.7258
Hungary	0.5525	0.7333
India	0.4673	0.6913
Italy	0.5988	0.7572
Japan	0.6211	0.7690
Korean Rep.	0.5525	0.7333
Mexico	0.3279	0.6276
Netherlands	0.5405	0.7272
Nigeria	0.3546	0.6393
Norway	0.6211	0.7690
Poland	0.5319	0.7229
Portugal	0.5000	0.7071
Russian Fed.	0.3731	0.6476
S. Africa	0.2809	0.6075
Slovakia	0.6806	0.8014
Spain	0.5405	0.7272
Sweden	0.6289	0.7732
Switzerland	0.5376	0.7258
USA	0.4673	0.6913
Venezuela	0.3788	0.6501

4.5 Dominance

Median curves Lorenz-dominate mean value curves for identical parameters. This means that

$$F_\varepsilon(x) \geq G_\varepsilon(x)$$

for all $x \in [0, 1]$ and for all common parameters $\varepsilon > 0.5$; for Lorenz-dominance see figure 2. Whenever the fractions of the median and the mean value are chosen equally, the induced distribution of the median is more uneven than that of the mean value. In addition, Lorenz-dominance extends to the case of median parameters being smaller than equity parameters ($\varepsilon \geq \eta > 0.5$)

$$F_\varepsilon(x) \geq G_\eta(x)$$

for all $x \in [0,1]$ but also to the more interesting case of certain larger median parameters. The latter situation is described by the inequality

$$F_\varepsilon(x) = G_{\eta(\varepsilon)}(x) \geq G_\eta(x)$$

which is valied for all $\eta \leq \eta(\varepsilon)$.

5 Summary and conclusion

The robustness issue of the median is advocated to be a non-issue in the context of defining a poverty line. The argument is that the median undoubtedly is robust but robust "around an inappropriate value". In contrast to the mean value, the median conceptually limits the number of poor to 50% of the population when the poverty line is less than the average income. The latter however, seems widely accepted.

At best, the median may be suited for defining a poverty line in rather balanced income groups. But it seems to be an improbable task to separate the income distributions that might admit a median-based poverty line from those that certainly do not. In particular, income distributions from countries in transition to more overall income and development, as well as those in tansition to less overall income and less development would be difficult to assign to either type. Globalization motivates the quest for a common poverty notion for nations and regions (aggregations of nations). Thus, mean value-based poverty lines are superior to median-based poverty lines.

The difference of using either median or mean value for defining poverty becomes obvious when extrapolated to the whole income distribution. From the view points of fairness and mere interest, not only the poorest should be compared to all others but each individual should be given the perspective of being the poorest. The equity parameter and the median parameter give each individual a hint on that perspective. The two parameters are nothing but approximate fractions of mean value and median of all larger incomes. A level of inequality which, from this perspective, is acceptable by many should provide for cohesion within an income group. But extending the perspective of 50% or 60% of the median to all results in income distributions that are at least as unbalanced as the Brazilian distribution. Extending the perspective of 50% or 60% of the mean value to all individuals results in income distributions that are similar to many European distributions which constitute reasonable prototype distributions for the globe.

As a conclusion from the foregoing, the mean value instead of the median is seen to be the reasonable and uniformly applicable notion of averge income that allows to specify a relative poverty line. The issue of whether this should be 50% or 60% or else is secondary and any fixed value in the range of 50% to 60% is plausible.

Considering all arguments, it can only be speculated about the true reasons for the median having swept away the mean value for poverty analysis in official statistics. The spectrum of these reasons ranges from an accidental decision to camouflage of the state of affairs in which the world lives. All those reasons should be laid open and debated – if it were only for the purpose of proving the soundness of pertinent statistics, the absence of any deliberate bias and the absence of any levels of inequality and poverty that are covertly tolerated.

References

[BeLi] Bellù, L.G., Liberati,, P., "Impacts of Policies on Poverty, Axioms for Poverty Measurement", Food and Agriculture Organization of the United Nations, FAO, report, Rome, 2005, www.fao.org/tc/easypol.

[Bu1] Bundesregierung, "Lebenslagen in Deutschland, Der erste Armuts- und Reichtumsbericht der Bundesregierung", Berlin, 2001, www.bmas.de/coremedia/generator/902/property=pdf/lebenslagen __in__deutschland__armutsbericht1.pdf.

[Bu2] Bundesregierung, "Lebenslagen in Deutschland, Der dritte Armuts- und Reichtumsbericht der Bundesregierung", Berlin, 2008, www.bmas.de/coremedia/generator/26742/property=pdf/dritter__armuts__und__reichtumsbericht.pdf.

[Ea] Easton, B., "Beware the median", 2002, www.eastonbh.ac.nz/?p=28.

[ER] Esteban, J., Ray, D., "On the measurement of polarization, Econometrica 62, 1994, p. 819-851.

[K] Kämpke, T., "Efficient quantization algorithms for discrete functions", Journal of Interdisciplinary Mathematics 6, 2003, p. 83-115.

[Kas] Kaskosz, B., "3D function grapher", University of Rhode Island, Kingston, 2004, www.math.uri.edu/~bkaskosz/flashmo/g3dblack.

[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., "Equity analysis by functional approach", in: D. Baier et al. (eds.), "Data Analysis and Decision Support", Springer, Berlin, 2005, p. 241-248.

[Le] Lerman, R.I., "Globalization and the Fight Against Poverty", Urban Institute, Washington, 2002, www.urban.org/url.cfm?ID=410612.

[LoRo] Lopuhaä, H.P., Rousseeuw, P.J., "Breakdown points of affine equivariant estimators of multivariate location and covariance matrices", The Annals of Statistics 19, 1991, p. 229-248.

[Ma] Mandel, S., "Debt relief as if people mattered", new economics foundation nef, London, 2006, www.jubileeresearch.org/news/debt_relief_final.pdf.

[MSt] Müller, A., Stoyan, D., "Comparison methods for stochastic models and risks", Wiley, New York, 2002.

[NaDo] Nakaya, T., Dorling, D., "Geographical inequalities of mortality by income in two developed island countries: a cross-national comparison of Britain and Japan", Social Science and Medicine 60, Issue 12, 2005, p. 2865-2875.

[No] Noll, H.-H., " 'At Risk of Poverty Rate': The flagship among the Laeken Indicators? - a note of caution", Luxembourg Presidency Conference on Taking forward the EU Social Inclusion Process, Luxembourg, 2005.

[Rey] Reynolds, A., "Mean vs. median", Working paper, Cato Institute, Washington, 2007, www.cato-unbound.org/2007/02/24/alan-reynolds/mean-vs-median-income.

[Rod] Rodriguez, J.G., "Measuring polarization, inequality, welfare and poverty", Economic Working Paper no. 75, Centro de Estudios Andaluces, Sevilla, 2004.

[SaSm] Saunders, P., Smedding, T., "Beware the mean!", Social Policy Research Center, University of New South Wales, Sydney, Newsletter 81, May 2002, p. 1 ff.

[Sh] Shanley, W., "Poverty in America: American dream now a nightmare for millions", 2007, www.globalresearch.ca/PrintArticle.php?articleId=5471.

[USC] US Census Bureau, "Income, poverty and health insurance coverage in the United States: 2007", Current population reports 60-235, US Government printing office, Washington, 2008.

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