MEASURING INEQUALITY

Giovanni Vecchi

Dept. of Economics
Universitá di Roma "Tor Vergata"
(giovanni.vecchi@uniroma2.it)

Serbia National Poverty Analysis Workshop
March 31-April 4, 2008

Plan of the Lecture

Functional *vs.* Personal Income Distribution

- Economists make a distinction between:
  - **Functional distribution of income**
    Distribution among factors of production (land, labor and capital)
  - **Personal** (or **size**) **distribution of income**
    Distribution among persons, irrespective of their economic function.
- In **this lecture** we will focus on the latter.

The Personal Distribution of Income

Table 1.1 *Distribution of incomes, UK, 1994/5*

Income range	Before tax*		After tax**	
	No. of incomes (000s) (1)	Total income (£m.) (2)	No. of incomes (000s) (3)	Total income (£m.) (4)
£3,445–£3,999	1,190	4,440	1,330	4,950
£4,000–£4,499	1,010	4,280	1,150	4,910
£4,500–£4,999	1,010	4,800	1,160	5,480
£5,000–£5,499	943	4,950	1,110	5,860
£5,500–£5,999	951	5,480	1,100	6,300
£6,000–£6,999	1,740	11,300	2,190	14,200
£7,000–£7,999	1,680	12,600	2,050	15,400
£8,000–£9,999	3,060	27,500	3,740	33,600
£10,000–£11,999	2,760	30,300	3,080	33,800
£12,000–£14,999	3,350	44,900	3,470	46,600
£15,000–£19,999	3,920	67,800	3,490	60,000
£20,000–£29,999	3,530	84,000	2,140	50,500
£30,000–£49,999	1,220	45,100	627	23,000
£50,000–£99,999	387	25,600	189	12,400
£100,000–£199,999	93	12,400	37	4,840
£200,000 and over	25	9,490	11	3,770

Notes: * By range of income before tax.
** By range of income after tax.
Source: Board of Inland Revenue, *Inland Revenue Statistics*, 1996, HMSO, London, table 3.3, page 35.

Income Inequality is Not Self-Defining

Focus on the term '**inequality**'.

- When we say **income inequality**, we mean simply **differences in income**, without regard to their desirability as a system of reward or undesirability as a scheme running counter to some ideal of equality.

  Kuznets (1953: xxvii)

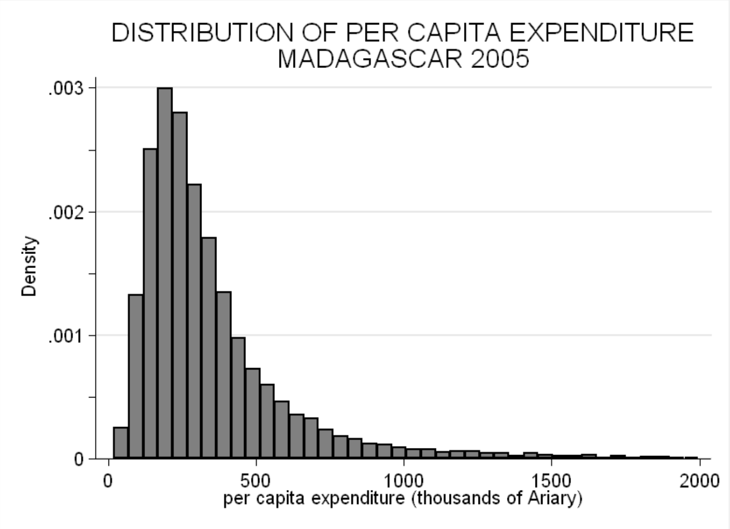
Q. **In practice**, how can we appraise the inequality of a given income distribution?

A. Three main options:

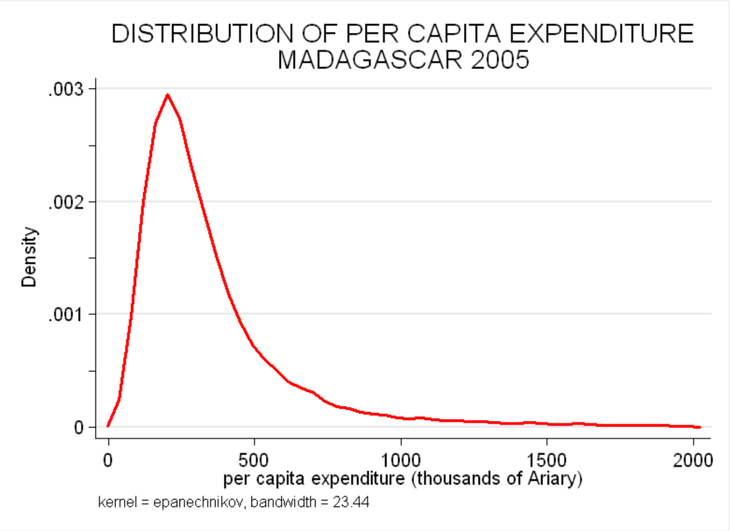
1. tables
2. graphs
3. summary statistics

The Histogram

A Familiar Graph



The Probability Density Function



The Probability Density Function

Definition

- The **cumulative distribution function (cdf)**:

$$F(x) = \int_{-\infty}^{x} f(t)dt$$

- If X is income and, say, $x = 2,000$€, then $F(x) = \Pr(X \leqslant 2,000)$, that is the fraction of people with less than 2,000 €.
- The **probability distribution function (pdf)** is the derivative of the cdf:

$$f(x) = \frac{dF(x)}{dx}$$

The Probability Density Function

Interpretation

- By definition of derivative:

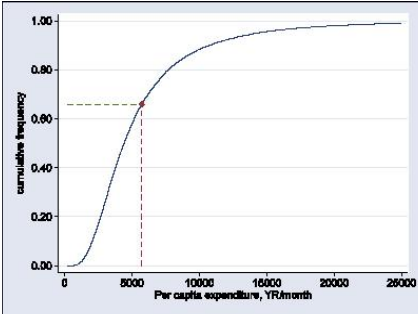
$$f(x) = \frac{dF(x)}{dx} = \underbrace{\lim_{h \to 0} \frac{F(x+h) - F(x)}{h}}_{\textit{difference quotient}}$$

- Now drop the limit (and replace $=$ by $\approx$):

$$\begin{aligned} f(x) \times h &\approx F(x+h) - F(x) \\ &\approx \Pr(x < X \leq x+h) \end{aligned}$$

- The **pdf** $f(x)$ is the probability of X falling in the interval $(x, x+h)$ divided by the length h of such an interval. It follows that **the pdf $f(x)$ is not a probability measure, but a scaled version of it**. In short, $f(x)$ may exceed one.

The Cumulative Density Function

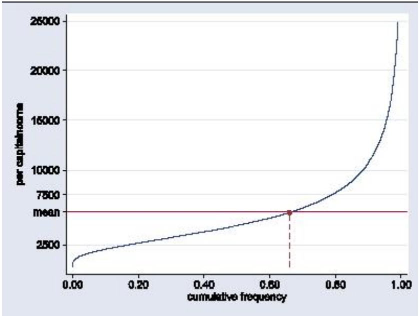


Pick up any income level on the x-axis, and the curve $F(x)$ will tell you the percentage of individuals in the population having a level of income lower than x.

The Parade of Dwarfs

Pen (1971)

- Assume that everyone in the population has **height** proportional to **income**.
- Line people up in order of height, and let them march.
- After some time, the shape of such a parade will be represented by the curve called **Parade of Dwarfs (and a Few Giants)**.



The Quantile Function

- Behind the Parade of Dwarfs, is the **quantile function**.
- Let $p = F(x)$ be the proportion of people in the population with income less than or equal to x. The **quantile function** $Q(p)$ is defined implicitly as

$$F\left[Q(p)\right] = p$$

or using the inverse cdf, as

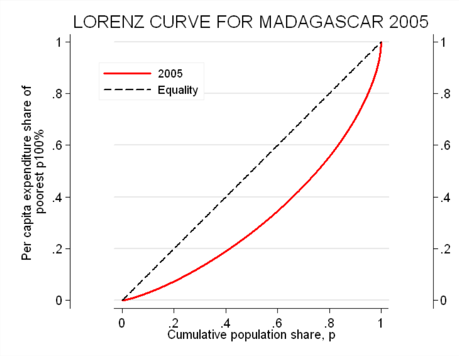
$$Q(p) = F^{-1}(p)$$

- *$Q(p)$ is the* **income level** *below which we find a proportion p of the population*.
Alternatively, it is the income of that individual whose rank - or percentile - in the distribution is p.

The Lorenz Curve (1905)

Picture & Intuition

- **Horizontal axis**: cumulative % of population (individuals ordered *poorest to the richest*)
- **Vertical axis**: cumulative % of income received by each cumulative % of population.
- **45-degree line**: Lorenz curve if perfect equality.
- The "overall" **distance** between the 45-degree line and the Lorenz curve is indicative of the amount of **inequality** present in the population.



The Lorenz Curve (1905)

Mathematically

- The Lorenz curve is defined as follows:

$$L(p) = \frac{\int_0^p Q(q)dq}{\int_0^1 Q(q)dq} = \frac{1}{\mu} \int_o^p Q(q)dq$$

- The **numerator** sums the incomes of the poorest p% of the population;
- The **denominator** sums the incomes of all. Since population size is normalized to 1, the denominator gives average income.
- The **ratio** L(p) thus indicates the cumulative % of total income held by a cumulative proportion p of the population.
- For **example**, if $L(\mathbf{0.5}) = 0.3$, then we know that the **50%** poorest individuals hold 30% of the total income in the population.

See Duclos and Araar (2006).

From Graphs to Summary Measures of Inequality

- Tables and charts are fine, but a better **conceptual understanding** comes from constructing inequality measures from **first principles**.
- We shall start from the most straightforward approach: inequality measures as **pure statistical measures of dispersion**.

Measures of Dispersion

1. **Range**

$$R = x_{\max} - x_{\min}$$

PRO Easy to compute and communicate.

CON Completely insensitive to changes in incomes between the extremes.

2. **Variance**

$$\sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

PRO Easy to compute, additively decomposable

CON Depends on the scale of measurement.

3. **Coefficient of Variation**

$$\kappa = \frac{\sqrt{\sigma^2}}{\bar{x}}$$

Quantiles, Quartiles, Quintiles, . . .

1. The p–**quantile** of a distribution of values **is a number** x_p such that a proportion p of the population values are less than or equal to x_p.
2. For example, if $p = 0.5$, then the 0.5–quantile $x_{0.5}$ is any value such that

$$F(X < x_{0.5}) = 0.5$$

3. Certain quantiles have **special names**:
   - The 0.5–quantile $x_{0.5}$ is the **median**, or **50-th percentile**.
   - The 0.1–quantile is the **first decile**, or **10-th percentile**.
   - The 0.2–quantile is the **first quintile**, or **20-th percentile**.
   - The 0.25–quantile is the **first quartile** Q_1, or **25-th percentile**.
   - etc. etc.
4. Quantiles are used to calculate the **interquartile range**

$$IQR = Q_3 - Q_1$$

and other measures of spread . . .

Quantile Ratios

Definition

- A **quantile ratio** measures the **gap** between the rich and the poor.
- It is defined as the **ratio** of two **quantiles**, $Q(p_2)/Q(p_1)$ using percentiles p_1 and p_2.
- Two popular indices are:
  1. the **quartile ratio** ($p_2 = 0.75$ and $p_1 = 0.25$):

$$QR = \frac{Q(p_{75})}{Q(p_{25})}$$

  2. the **decile ratio** ($p_2 = 0.90$ and $p_1 = 0.10$):

$$DR = \frac{Q(p_{90})}{Q(p_{10})}$$

Quantile *Share* Ratios

(Please note the red italics in the title.)

- Let S_{20} denote the **share** of (equivalised disposable) income received by the *bottom* 20 % of the population, and S_{80} the income share received by the *top* 20% of the population.
- The **quintile share ratio** is defined as follows:

$$S_{80-20} = \frac{S_{80}}{S_{20}}$$

- The quintile share ratio is *the* level-1 Laeken indicator, chosen by the EU to monitor income distribution.
- The **EU25** average was **4.9** in 2005, which means that the wealthiest quintile had 4.9 times more income than the poorest. Ratios range from **3.3** (Sweden) to **8.2** (Portugal).

The Gini Coefficient

1. Yitzhaki (1997) counts more than a dozen formulas available for the Gini index.
2. The classical definition of the **Gini coefficient** is probably as follows:
$$G = \frac{1}{2n^2\bar{x}} \sum_{i=1}^{n} \sum_{j=1}^{n} \mid x_i - x_j \mid$$
3. The Gini coefficient ranges from 0 (all recipients have the same income: full equality), to 100 (all income is received by one recipient: maximum inequality).
4. The Gini index for **Serbia** in 2006 was 0.28.

Summing Up

Q1. Would all summary measures lead to similar results?

A1. **No!**

Q2. Does the use of such an apparently neutral statistic conceal judgements about the desirability of different forms of redistribution?

A2. **Yes!**

- The problem is that is hard to be aware of value judgments unless a more **theoretical approach** is considered.
- Atkinson (1970) pioneered the use of **social welfare functions**.

Inequality and Social Welfare

1. **Social Welfare Functions** (SWF) provide the basis for
   - making inequality judgments
   - deriving inequality measures consistent with judgments.
2. Time does not permit to cover this topic properly, but I will describe the main steps of the approach.

Social Welfare Functions

1. Take a ‘regular’ **SWF**

$$W(x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n} U(x_i)$$

2. Assume a concave **social marginal utility**

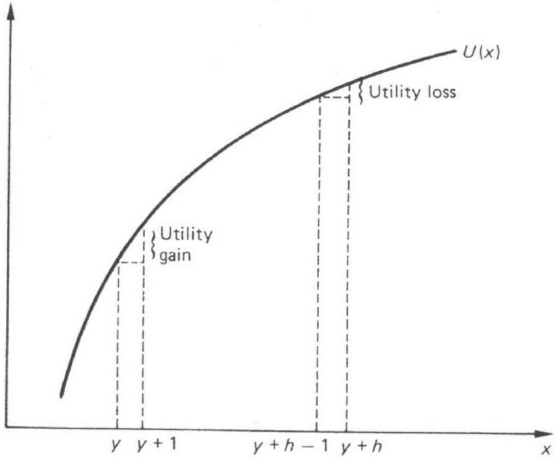
$$U(x_i) = \frac{x^{1-\varepsilon} - 1}{1 - \varepsilon}$$

where $U(x_i)$ has constant elasticity ε, **inequality aversion parameter**.

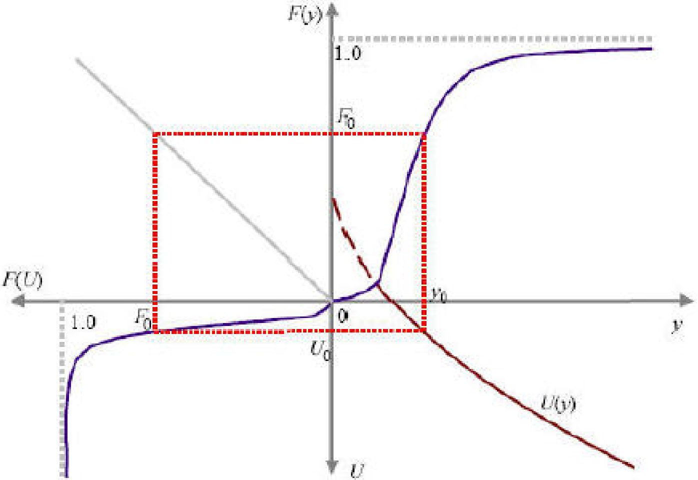
3. Obtain $F(U)$ from x, via $F(x)$
4. Work out $f(U)$, which is simply the slope of $F(U)$

The Concavity of Social Marginal Utility

Interpretation



The Income Distribution and the Distribution of Social Utility



The Atkinson's Inequality Index

Definition

1. The **Atkinson's inequality index** (for inequality aversion ε)

$$A_\varepsilon = 1 - \frac{x_e}{\bar{x}} = 1 - \left[\frac{1}{n} \sum_{i=1}^{n} \left[\frac{y_i}{\bar{x}} \right]^{1-\varepsilon} \right]^{\frac{1}{1-\varepsilon}}$$

2. The key point is the choice of the value of ε:
   - **low** $\varepsilon =$ we are sensitive to changes among top incomes
   - **high** $\varepsilon =$ we are concerned mostly with low incomes

Inequality and Information Theory

- Inequality measures can be derived from an analogy with **information theory**.
- The key concept: **entropy**, the degree of disorder of a system.
- Henri Theil (1967) obtained a measure of inequality, the **Theil's entropy index**:

$$T = \sum_{i=1}^{n} s_i \left[\log s_i - \log \left(\frac{1}{n} \right) \right]$$

  where s_i is the share of the i-th person in total income.
- The index T is a member of a wider class of inequality measures, the **generalized entropy inequality** (GEI) class.

Deriving Inequality Measures From Axioms

- **Axiom**: a statement accepted as true as the basis for argument or inference.[1]
- The **axiomatic approach** allows us to obtain a mathematical formula that delivers a class of inequality measures that satisfy a set of elementary properties (axioms) that we think inequality measures ought to have.

[1] Merriam-Webster

Five Axioms of Inequality Measures – 1/4

(A) Anonimity (or Symmetry) – If X is any permutation of Y, then $I(X) = I(Y)$.

- *In short, it does not matter who is earning the income.*

(P) The Population Principle – When one income distribution is an n-fold replication of another, the two distributions are distributionally equivalent.

- *The population size does not matter: all that matters are the proportions of the population who earn different levels of income.*

Five Axioms of Inequality Measures – 2/4

(S) The Scale Invariance (or Relative Income Principle) – If everyones income changes by the same proportion, then inequality does not change.

$$X = (x_1, x_2, \ldots, x_n)$$
$$Y = (\lambda x_1, \lambda x_2, \ldots, \lambda x_n)$$
$$\Rightarrow I(X) = I(Y)$$

- *Inequality should not depend on whether income is measured in USD or €.*
- *Income* **levels**, *in and of themselves, have no meaning as far as* **inequality measurement** *is concerned.*

Five Axioms of Inequality Measures – 3/4

(T) The (Pigou-Dalton) Principle of Transfers – If one distribution is obtained from another by transferring a positive amount δ of income **from** a relatively **rich** person **to** a relatively **poor** person without altering their ranks in the distribution, then **inequality must decrease**.

$$X = (x_1, x_i, \ldots, x_j, \ldots, x_n)$$
$$Y = (x_1, x_i + \delta, \ldots, x_j - \delta, \ldots, x_n), \qquad \delta > 0$$
$$\Rightarrow I(Y) \leq I(X)$$

Five Axioms of Inequality Measures – 4/4

(D) Decomposability (or subgroup consistency) – An **additively decomposable inequality measure** is one which can be expressed as a weighted sum of the inequality values calculated for population groups plus the contribution arising from differences between group means

$$I = \sum_{k=1}^{K} \omega_k I_k + I(\bar{x}_1, \ldots, \bar{x}_k), \qquad \sum_{k=1}^{K} \omega_k = 1$$

where I_k is the inequality index calculated within the k-th group, and ω_k are population shares.

Shorrocks (1980) "Theorem 5"

Theorem

Any inequality measure that simultaneously satisfies the properties of anonimity, scale independence, population principle, the principle of transfers, and decomposability, must have the following form

$$GE(\theta) = \frac{1}{\theta^2 - \theta}\left[\frac{1}{n}\sum_{i=1}^{n}\left[\frac{x_i}{\bar{x}}\right]^{\theta} - 1\right]$$

or some ordinally equivalent transformation of it (θ is a real parameter).

- The class $GE(\theta)$ has become known as the **Generalized Entropy Indices**.

The Generalized Entropy Indices

- Depending on the value of the θ parameter:
  - $\theta = 0 \Rightarrow$ **Mean Logarithmic Deviation**

$$GE(0) = MLD = \frac{1}{n}\sum_{i=1}^{n} \log\left(\frac{\bar{x}}{x_i}\right)$$

  - $\theta = 1 \Rightarrow$ **Theil Index**

$$GE(1) = THEIL = \frac{1}{n}\sum_{i=1}^{n} \frac{x_i}{\bar{x}} \log\left(\frac{x_i}{\bar{x}}\right)$$

  - $\theta = 2 \Rightarrow$ Half **Coefficient of Variation Squared**

$$GE(2) = \frac{\kappa^2}{2} = \frac{1}{n}\sum_{i=1}^{n} \frac{x_i}{\bar{x}} \log\left(\frac{x_i}{\bar{x}}\right)$$

The Structure of Inequality

- **Inequality decompositions** are typically used to estimate the extent to which the heterogeneity of the population affects overall inequality.
- Two popular techniques are:
  1. Decomposition **by population sub-group**
  2. Decomposition **by income source**
- In this lecture, we will focus on the former.

Decomposition by Population Sub-Group

Shorrocks (1980)

- Societies can often be partitioned into **groups** (North-South, age brackets, gender, ethnic groups, etc.).
- We would like to be able to write a formula giving **total** inequality as a function of inequality **within** the constituent groups, and inequality **between** the groups:

$$I_{TOTAL} = I_{WITHIN} + I_{BETWEEN}$$

Shorrocks (1980)

Theory

1. The most popular additively decomposable inequality index is the **Mean Logarithmic Deviation**.
2. Partition the population into $k = 1, \ldots, K$ groups. Then:

$$MLD = \underbrace{\sum_{k=1}^{K} v_k MLD_k}_{WITHIN} + \underbrace{\sum_{k=1}^{K} v_k \log\left(\frac{\bar{x}}{\bar{x}_k}\right)}_{BETWEEN}$$

where v_k are population shares.

Shorrocks (1980)

Example

Table 5 – Inequality Decompositions, Madagascar 2001 and 2005

	2001			2005		
	Index	**Within**	**Between**	**Index**	**Within**	**Between**
by urban/rural area						
MLD	0.371	0.332	0.040	0.224	0.209	0.015
(%)	*(100)*	*(89)*	*(11)*	*(100)*	*(93)*	*(7)*
Theil	0.402	0.359	0.043	0.273	0.258	0.015
(%)	*(100)*	*(89)*	*(11)*	*(100)*	*(95)*	*(5)*
by province						
MLD	0.371	0.306	0.065	0.224	0.213	0.010
(%)	*(100)*	*(83)*	*(17)*	*(100)*	*(95)*	*(5)*
Theil	0.402	0.335	0.067	0.273	0.263	0.010
(%)	*(100)*	*(83)*	*(17)*	*(100)*	*(96)*	*(4)*

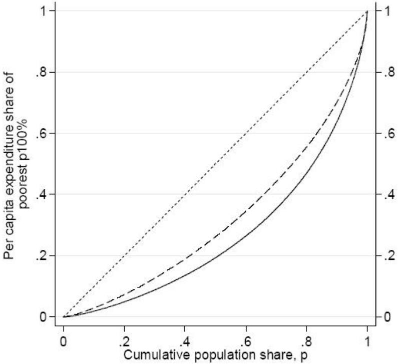
Source: Amendola and Vecchi (2007)

Inequality Comparisons

- Inequality measures are most frequently used for **dynamic comparisons** (comparing inequality measures across time), and for **policy analysis** (*e.g.* to compare inequality across regions or by population sub-groups).
- Despite their seeming straightforwardness, inequality comparisons are **tricky**.
- One may consider comparing distributions:
  1. by means of **summary measures** such as the Gini coefficient (cardinal approach)
  2. by means of **Lorenz curves** (ordinal approach)

Lorenz Dominance

Definition



- If the Lorenz curve of distribution X_A lies nowhere below and at least somewhere above the Lorenz curve of distribution X_B then X_A Lorenz-dominates X_B:

$$L_A > L_B$$

- When one distribution Lorenz-dominates another, then the one that dominates is more equal than the one that is dominated.

The Connection Between Lorenz Curves and Inequality Axioms

- Any inequality measure is consistent with the Lorenz dominance if and only if it is simultaneously consistent with the anonimity, population, scale invariance and Pigou-Dalton principles.
- In other words, all inequality measures satisfying the 4 axioms above will agree in their **inequality ranking** and will rank distributions consistently with the Lorenz criterion.

The Connection Between Lorenz Dominance and Social Welfare Functions

Theorem (Atkinson 1970)

Let X_A and X_B denote two income distributions with the same mean $\bar{x}_A = \bar{x}_B$. Let W be a 'regular' SWF. Then:

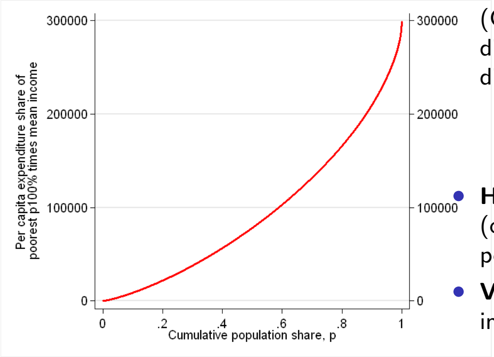
$$W_A > W_B \Leftrightarrow L_A > L_B$$

- When the means are equal, an income distribution that Lorenz-dominates another is not only more equal but **better** in welfare terms (provided we accept the properties of W).

Unfortunately, Lorenz Curves Can Cross . . .

- Lorenz domination gives a **partial ordering** of distributions: when Lorenz curves cross, neither distribution dominates the other.
- Moreover, because Lorenz curves are **unaffected by the mean** of the distribution, they cannot be used to rank distributions with different mean income.

The Generalized Lorenz Curve (Shorrocks, 1983)



- **Generalized Lorenz Curves** (GLC) are used for comparing different distributions with different means.

$$GL(p) = \bar{x}L(p) = \int_0^p Q(q)dq$$

- **Horizontal axis**: same as Lorenz (cumulative % of population, poorest to the richest)
- **Vertical axis**: cumulative share of income <u>times</u> mean income.

Generalized Lorenz Dominance

Theorem (Shorrocks 1983)

Let X_A and X_B denote any pair of income distributions. Let W be a 'regular' SWF. Then:

$$W_A \geq W_B \Leftrightarrow GL_A \geq GL_B$$

- If the GLCs for two distributions do not intersect, then we can say that the distribution with the higher curve is the socially preferred one, according to all 'regular' SWFs.
- If two GLC intersect the theorem does not apply. Thus, the GLC dominance check also yields an incomplete ranking (like the ordinary Lorenz dominance).

Summary & Conclusions

1. **Small-font footnotes** and technical appendixes matter.
2. In practice, '**the way that data come**' dictates many of the definitional issues.
3. There is **no single best measure** of income inequality.
4. Inequality measures are **not purely statistical**. The use of a particular inequality measure embodies a **social judgement** about the weight to be attached to the inequality at different points on the income scale. Judgements can either be implicit (*e.g.*, Gini Coefficient) or explicit (*e.g.*, Atkinson Index).

Summary & Conclusions

I can take any country and prove that in some period (whatever it is) inequality has increased or decreased in it, or any two countries and prove that inequality is higher in the one or in the other, by choosing different inequality measures, all of which would probably seem good and valuable at first sight.

Kolm (1976: 416)

Summary & Conclusions

1. **Lorenz Curves** provides a powerful device for ranking distributions (with the same mean) from the welfare point of view.
2. If **Lorenz Curves** intersect it is no possible to rank distributions.
3. **Generalized Lorenz Curves** can be used to rank distributions with different mean income and/or with intersecting Lorenz Curves.

For Further Reading I

F. Cowell (1992)
Measuring Inequality, 2nd edn.
Hemel Hempstead: Harvester Wheatsheaf.

J.-Y. Duclos and A. Abdelkrim (2006)
Poverty and Equity: Measurement, Policy and Estimation with DAD
New York: Springer.

Downloadable from: http://www.idrc.ca/en/ev-97152-201-1-DO_TOPIC.html

A. Deaton (1997)
The Analysis of Household Surveys.
Baltimore-London: John Hopkins University Press.

A.B. Atkinson (1970)
On the Measurement of Inequality.
Journal of Economic Theory, 2: 244–73.

For Further Reading II

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

A.F. Shorrocks (1982)
Inequality Decomposition by Factor Components
Econometrica, 50(1): 193-211.

A.F. Shorrocks (1982)
The Impact of Income Components on the Distribution of Family Incomes
Quarterly Journal of Economics, 98(2): 311-326.

For Further Reading III

D. Mookherjee and A.F. Shorrocks (1982)
A Decomposition Analysis of the Trend in UK Income Inequality
Economic Journal, 92: 886-902.

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

S.P. Jenkins (1995)
Accounting for inequality trends: Decomposition analyses for the UK, 1971-86
Economica, 62: 29-63.