MPRA
Munich Personal RePEc Archive

Charting Income Inequality: The Lorenz Curve

Lorenzo G. Bellù and Paolo Liberati

Food and Agriculture Organization of the United Nations

1. November 2005

Online at https://mpra.ub.uni-muenchen.de/30063/
MPRA Paper No. 30063, posted 7. April 2011 18:46 UTC

EASYPol

Resources for policy making

ANALYTICAL TOOLS

Module 000

Charting Income Inequality

The Lorenz Curve

EASYPol

Resources for policy making

Charting Income Inequality

The Lorenz Curve

by

Lorenzo Giovanni Bellù, Agricultural Policy Support Service, Policy Assistance Division, FAO, Rome, Italy

Paolo Liberati, University of Urbino, "Carlo Bo", Institute of Economics, Urbino, Italy

for the

Food and Agriculture Organization of the United Nations, FAO

About EASYPol

The EASYPol home page is available at: *www.fao.org/easypol*

EASYPol is a multilingual repository of freely downloadable resources for policy making in agriculture, rural development and food security. The resources are the results of research and field work by policy experts at FAO. The site is maintained by FAO's Policy Assistance Support Service, Policy and Programme Development Support Division, FAO.

This modules is part of the resource package Analysis and monitoring of socio-economic impacts of policies.

The designations employed and the presentation of the material in this information product do not imply the expression of any opinion whatsoever on the part of the Food and Agriculture Organization of the United Nations concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries.

© FAO November 2005: All rights reserved. Reproduction and dissemination of material contained on FAO's Web site for educational or other non-commercial purposes are authorized without any prior written permission from the copyright holders provided the source is fully acknowledged. Reproduction of material for resale or other commercial purposes is prohibited without the written permission of the copyright holders. Applications for such permission should be addressed to: **copyright@fao.org**.

Table of Contents

1 SUMMARY

This module explains how to build **Lorenz Curves** for income distributions and discusses their use for inequality measurement. A short conceptual background, a step-by-step procedure and a simple numerical example illustrate how to calculate and draw Lorenz Curves. A discussion on the use of Lorenz Curves to represent inequality is also provided. It highlights that the Lorenz Curve is one of the most used ways of representing income distributions in empirical works thanks to its immediate comparability with a "natural" benchmark, the **Equidistribution line**, representing the most egalitarian distribution. The concepts of **Lorenz dominance** and **intersection of Lorenz Curves** are also discussed. Furthermore, the appendix provides a detailed presentation of the properties of the Lorenz Curves.

2 INTRODUCTION

Objectives

The aim of this module is to illustrate the functioning and the meaning of one of the most popular tools to measure inequality in an income distribution, i.e., the Lorenz Curve. The idea being to develop methods to allow analysts to build Lorenz Curves starting from any income distribution and, possibly, to compare inequality among different income distributions.

Target audience

This module targets all policy analysts who work in either public administrations (central and local), NGOs, political parties, professional organizations or in consulting firms that are willing to enhance their expertise in analyzing inequality in income distributions. Lecturers in selected undergraduate courses in the economics of inequality and poverty (and other related fields) may also be interested in using this material for academic purposes.

Required background

The module should be delivered to an audience that is already familiar with basic mathematics and statistics. In addition, the concept of income distributions and **quantiles of incomes** should be covered before delivering this module.

To find relevant materials in these areas, the reader can follow the links included in the text to other EASYPol modules or references[1]. See also the list of useful EASYPol links included at the end of this module.

[1] EASYPol hyperlinks are shown in blue, as follows:
a) training paths are shown in **underlined bold font**
b) other EASYPol modules or complementary EASYPol materials are in ***bold underlined italics***;
c) links to the glossary are in **bold**; and
d) external links are in *italics*.

3 CONCEPTUAL BACKGROUND

The Lorenz Curve is a tool used to represent income distributions as proposed by Lorenz (1905); it tells us which proportion of total income is in the hands of a given percentage of population[2]. This method is conceptually very similar to the method by quantiles. However, instead of ending up with income shares, the Lorenz Curve relates the cumulative proportion of income to the cumulative proportion of individuals.

The Lorenz Curve is obtained as follows:

The *x*-axis records the cumulative proportion of population ranked by income level. Its range is therefore (0,1).

The *y*-axis records the cumulative proportion of income for a given proportion of population, i.e. the income share calculated by taking the cumulated income of a given share of the population, divided by the total income Y, as follows:

$$L\left(\frac{k}{P}\right) = \frac{\sum_{i=1}^{k} y_i}{Y}$$

where :
$k=1....n$ is the position of each individual in the income distribution;
$i=1....k$ is the position of each individual in the income distribution;

P is the total number of individuals in the distribution;
y_i is the income of the i^{th} individual in the distribution
$\sum_{i=1}^{k} y_i$ is the cumulated income up to the k^{th} individual.

It is apparent that $\sum_{i=1}^{k} y_i$ ranges between 0, for k=0, and Y, for k=n, therefore $L\left(\frac{k}{P}\right) = \frac{\sum_{i=1}^{k} y_i}{Y}$ ranges between 0 and 1[3].

---

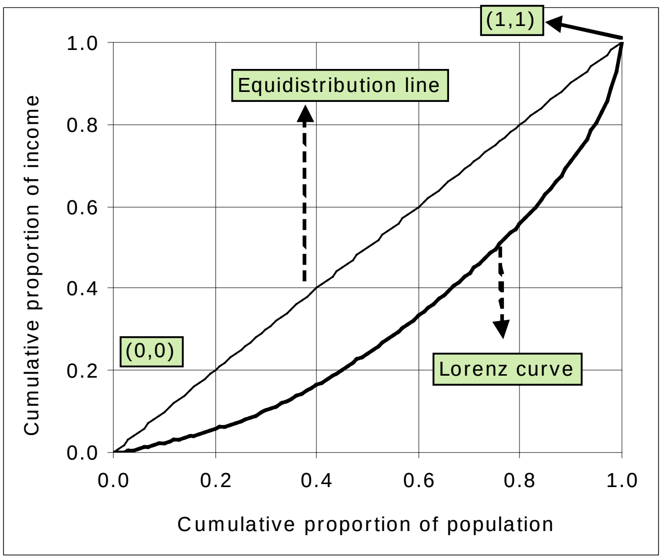
[2] Remember that an income distribution of a finite population of n individuals is an ordered list of incomes (from the lowest to the highest) where each income y_i is attached to a given individual or household i. The analytical representation is: $\mathbf{y} = (y_1, y_2,..., y_n)$ where $\mathbf{y}$ indicates a «vector» of individual incomes. If household incomes are considered, then to each household income there should also be attached a number w reflecting household size, in order to make meaningful comparisons among income levels: $\mathbf{y} = ((w_1, y_1),(w_2, y_2),...,(w_n, y_n))$. The case where incomes are owned by individuals is indeed a special case where w_i=1 for each observation.

[3] For a detailed analysis of the properties of the Lorenz curve, see the discussion in section 5, below, and the Appendix (section 8) to this module.

Figure 1, below, illustrates the shape of a typical Lorenz Curve. As you can see, in the graph, the curve starts from coordinates (0,0), as a zero fraction of the population owns a zero fraction of income. Since the Lorenz Curve records cumulative proportions, it must be that total population owns total income. Hence, the Lorenz Curve has coordinates (1,1) at the end point. Now, if income was equidistributed across a population (i.e. everybody had the same income), it would follow that a given proportion of the population (say 10 per cent) would have the same proportion of income (10 per cent). For example, if in a population of 100 individuals everybody had (1/100) of total income, the graph of the underlying income distribution would be the equidistribution line. Typically, an income distribution is made of poor and rich people. It means that poor individuals own less than an equidistributed share of total income because richer individuals own more than the same equidistributed share. Therefore, for a typical income distribution, the Lorenz Curve is the convex curve as depicted in Figure 1.

In the next sections a step by step procedure to work out a Lorenz Curve, a simple numerical example and a discussion on how to benckmark Lorenz Curves against the equidistribution line are presented.

Figure 1 - The Lorenz Curve and the equidistribution line



4 A STEP-BY-STEP PROCEDURE TO BUILD THE LORENZ CURVE

4.1 The step by step procedures

Figure 2, below, illustrates the path to build the Lorenz Curve.

To build a Lorenz Curve you first need to calculate the proportion of income belonging to each individual and which proportion of total population that same individual represents, Step 2. The next step is to cumulate both proportions, Step 3. Step 4, instead, defines the benchmark for the Lorenz Curve, the equidistributed line. The equidistribution line is defined on the basis of the most equalitarian income distribution:

each individual owns the same income, i.e., income is perfectly distributed among a given population[4]. In an ordinary income distribution, not all incomes are equal, which means that the Lorenz Curve will always lie below the equidistribution line. In Step 5, the two cumulative proportions are plotted.

Figure 2 - A step-by-step procedure to build Lorenz Curves

STEP	OPERATIONAL CONTENT
1	If not already sorted, sort the income distribution by income level
2	Define the proportion of income owned by each individual and his proportion on total population
3	Define the cumulative proportion of income and the cumulative proportion of population
4	Define the line of equidistributed income
5	Plot the cumulative proportion of income against the cumulative proportion of population

Remark: On this line proportion of income and proportion of population are the same.

4.2 An example of how to build Lorenz Curves

An example of a simple income distribution is given below in Table 1, below, where a list of 5 incomes is given, expressed, say, in US$/year.

Table 1 - An example of income distribution, US$/year

Individual	Income
1	2,417
2	7,800
3	8,489
4	10,072
5	12,957

In this income distribution, individual 1 owns US$2,417/year (he/she is the poorest), while individual 5 owns US$12,957/year (he/she is the richest). Table 2 below, illustrates the process to build the Lorenz Curve[5]:

[4] In the equidistribution line, the cumulative proportion of income is equal to the cumulative proportion of population (i.e., ten per cent of population owns ten per cent of income, twenty per cent of population owns twenty per cent of income, and so on).

[5] Most of the content of this tool will be discussed assuming «individuals» as reference units. Things do not change when «households» are considered. The adjustments needed to work with households are dealt with in the EASYPol Module 032: ***Equivalence Scales: General Aspects***, and related EASYPol Modules.

Table 2 - Calculating Lorenz Curves

Step 1		Step 2		Step 3		Step 4	Step 5
Order the income distribution		Define the proportion of income owned by each individual and the proportion of population corresponding to him		Define the cumulative proportion of income and population		Define the equidistribution line	Plot the cumulativ proportion of incor against the cumulat proportion of popula
Individuals	Income	Proportion of income belonging to each individual	Proportion of each individual on total population	Cumulative proportion of income	Cumulative proportion of population		See Graph below
1	2,417	0.058	0.200	0.058	0.200		
2	7,800	0.187	0.200	0.245	0.400		
3	8,489	0.203	0.200	0.448	0.600		
4	10,072	0.241	0.200	0.690	0.800		
5	12,957	0.310	0.200	1.000	1.000		
Total	41,735	1.000	1.000				

Remark: The equidistribution line can be built following Step 1 to 3 assuming that **everybody has the same income**

Step 1		Step 2		Step 3		Step 5
Order the income distribution		Define the proportion of income owned by each individual and the proportion of population corresponding to him		Define the cumulative proportion of income and population		Plot the cumulativ proportion of incor against the cumulat proportion of popula
Individuals	Income	Proportion of income belonging to each individual	Proportion of each individual on total population	Cumulative proportion of income	Cumulative proportion of population	See Graph below
1	8,347	0.200	0.200	0.200	0.200	
2	8,347	0.200	0.200	0.400	0.400	
3	8,347	0.200	0.200	0.600	0.600	
4	8,347	0.200	0.200	0.800	0.800	
5	8,347	0.200	0.200	1.000	1.000	
Total	41,735	1.000	1.000			

5 DISCUSSION

5.1 The Lorenz Curve to describe inequality

The Lorenz Curve is a very useful way to calculate income inequality. Figure 3, below, shows the shape of Lorenz Curves in the case of the three income distributions **A**, **B** and **C**, with the same total income.

Example 1: Case **A** is that of an ordinary income distribution. As incomes are lined up in ascending order, the best thing that can happen is that any given person can pass by with an income that is at most equal to the income of the next person that passes by; usually it is lower, as in the case of distribution **A**. As the Lorenz Curve cumulates proportions of income, when the first 10 per cent of the population has gone away, in an ordinary income distribution you **must** have handed out less than 10 per cent of total income. This means that there will be another 10 per cent of the population that will pass by with more than 10 per cent of total income. The shape of the Lorenz Curve is therefore of a convex shape.[6]

In Lorenz Curve representations, less inequality means that you will get a less pronounced convexity.

Example 2: Let us consider what happens in the case of equal incomes (the case of distribution **B**). In this case, each person that passes by has an equal share of total income, as incomes are all equal. As the Lorenz Curve cumulates proportions of income, when the first, say, 10 per cent of the population has gone away, you **must** have handed out 10 per cent of total income, as everybody has the same income. If it were that only 5 per cent of the total income had gone when the first 10 per cent of the population had passed by, then there would be some point in the parade where another 10 per cent of the population would pass by with more than 10 per cent of total income. But this would contradict the hypothesis that all incomes are equal. The case of equal incomes is therefore a case where, as far as population goes, in the parade the same percentage of total income goes with it. This gives rise to a linear Lorenz Curve, the so-called «equidistribution line». Indeed, for distribution **B**, the Lorenz Curve is a straight line. In this case, there is the minimum concentration of income. More precisely, there is **zero concentration**, as everybody has the same income.

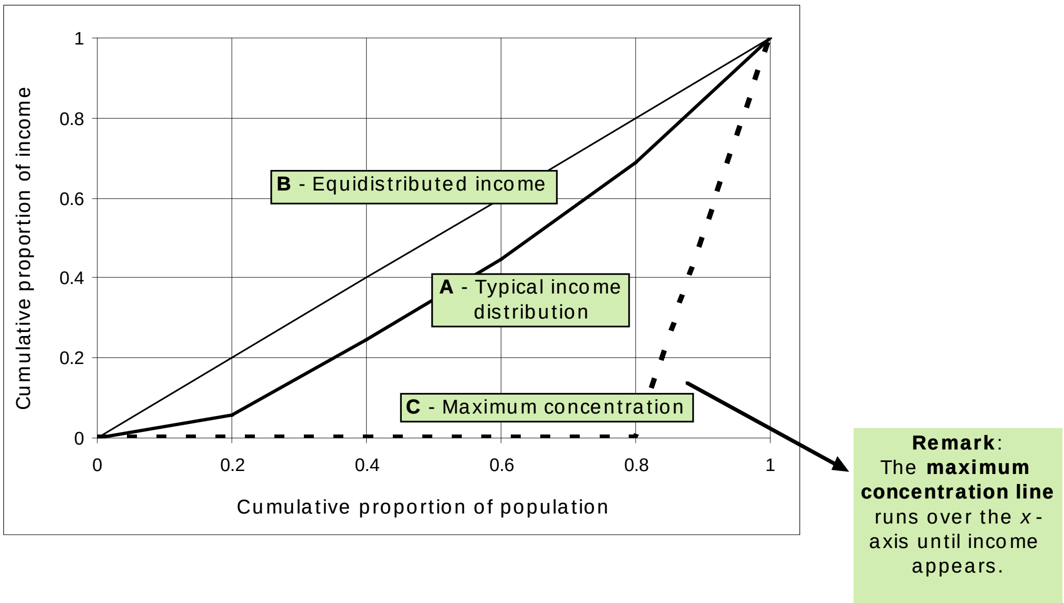
More inequality in income distributions implies more convex Lorenz Curves. At the extreme of distribution **C,** all persons except for one pass by with zero income, i.e., with zero per cent of total income. Only the last individual (or, say, the last 10 per cent of the population) will hold 100 per cent of total income. The Lorenz Curve is therefore a kinked curve, running on zero until the last individual is reached and then jumping to 100 per cent. In this specific case, there is the maximum concentration of income as possible.

The shape of the Lorenz Curve is therefore a good visual indicator of how much inequality there is in an income distribution. When incomes are less dispersed (i.e., there is little variability among incomes), the Lorenz Curve tend towards the equidistribution line. When incomes are more dispersed (i.e., there is a greater variability among incomes), the Lorenz Curve tend towards the kinked curve.

[6] See Appendix at the end of this module for a generalization of this property.

Figure 3 - Lorenz Curves and income distributions

Individuals	Income distribution A	Income distribution B	Income distribution C
1	2,417	8,347	0
2	7,800	8,347	0
3	8,489	8,347	0
4	10,072	8,347	0
5	12,957	8,347	41,735
Total	41,735	41,735	41,735



5.2 Lorenz dominance and intersection

Lorenz dominance of one income distribution over another occurs when, for any given cumulative proportion of population *p*, the Lorenz Curve of a given income distribution is above the Lorenz Curve(s) of the other distribution(s). Given the Lorenz Curve and its properties, the dominating Lorenz Curve implies an income distribution with less inequality.

However, there is no guarantee that given two income distributions one would Lorenz-dominate. It may be the case that Lorenz Curves **intersect**. In this case, by considering only Lorenz Curves, nothing can be said about which income distribution has less inequality.

It is worth noting that Lorenz dominance is based on a visual inspection of a particular way of representing income distributions. In this sense, it is not really a canonical index, as Lorenz dominance by itself cannot answer the question: «How much less» unequal is the income distribution[7]?

[7] We will see later that Lorenz dominance is an important concept, as it implies a particular class of inequality measures, the **Relative Inequality Indexes (RII)**.

All the properties of Lorenz dominance are the properties of the Lorenz Curve itself. There is only one important aspect of Lorenz dominance that is worth noting: **it is a partial ordering**. Lorenz Curves can indeed be used to rank income distributions according to their inequality, provided that they do not intersect. If they do, nothing can be said about income inequality.

Table 3, below, draws Lorenz Curves for three alternative income distributions. The first is the ordinary income distribution, income distribution **A**. The second and the third distributions (**X** and **Y** respectively) are obtained from the first by reducing the incomes of the poorest individuals and increasing incomes of the richest individuals, holding total income unchanged. In **Y**, income reductions and income increases are of greater intensity than in **X**.

Table 3 - Lorenz dominance

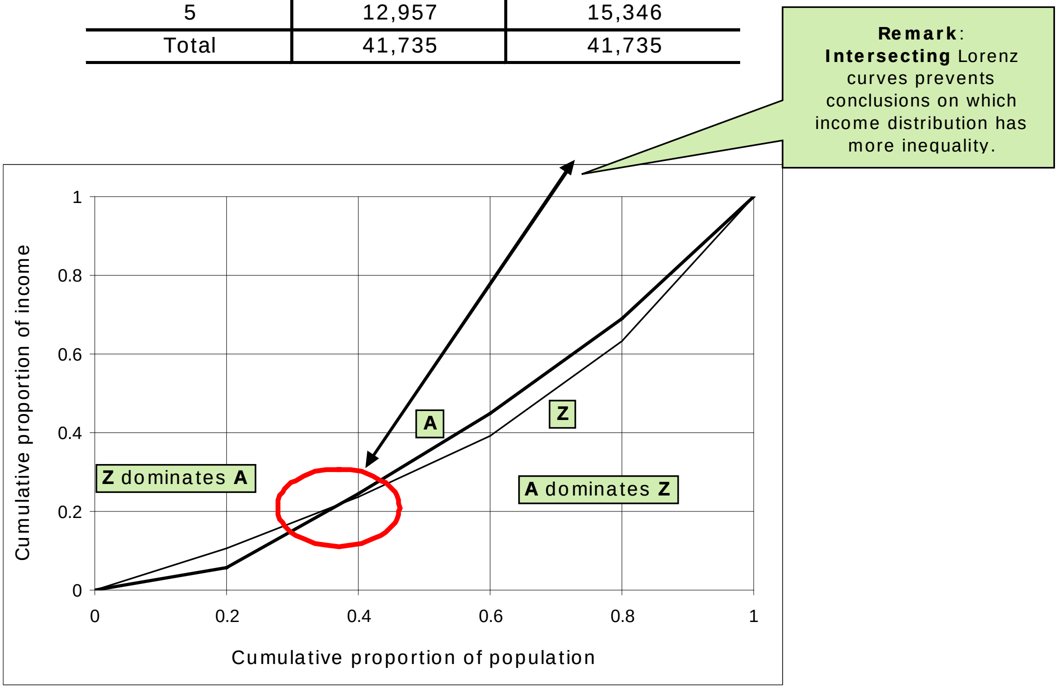
Individuals	Income distribution A	Income distribution X	Income distribution Y
1	2,417	417	117
2	7,800	7,800	2,800
3	8,489	8,489	8,489
4	10,072	10,072	10,072
5	12,957	14,957	20,257
Total	41,735	41,735	41,735

The graph associated to Table 3 depicts Lorenz Curves corresponding to each income distribution, obtained in the way described above. What is important in this context is the presence of Lorenz dominance. For example, the Lorenz Curve of income distribution **A** dominates both the Lorenz Curve of income distribution **X** and the Lorenz Curve of income distribution **Y**. In turn, the Lorenz Curve of income distribution **X** dominates the Lorenz Curve of income distribution **Y**. In all cases of Lorenz dominance, the dominating distribution has also less income inequality. This is the type of information you can get from Lorenz dominance.

However, it might be the case that Lorenz Curves intersect. Table 4, below, gives evidence of this possibility, by considering income distribution **A** and a transformation of it, **Z**, holding total income constant. As can easily be seen in the graph, there is a point where Lorenz Curves intersect. To the left of that point, the Lorenz Curve of income distribution **Z** dominates **A**. To the right of that point, the Lorenz Curve of income distribution **A** dominates **Z**. In this case, nothing can be said about which distribution has more income inequality.

Table 4 - Lorenz Curve intersection

Individuals	Income distribution A	Income distribution Z
1	2,417	4,417
2	7,800	5,400
3	8,489	6,500
4	10,072	10,072
5	12,957	15,346
Total	41,735	41,735



6 READERS' NOTES

6.1 Time requirements

Time required to deliver this module is estimated at about three hours.

6.2 EASYPol links

The Lorenz Curve has a strict link with a very popular inequality index, the **Gini Index**. The Gini index and its correspondence with the Lorenz Curve will be discussed in the EASYPol Module 040: ***Inequality Analysis: The Gini Index.***

Other useful links are:

EASYPol Module 008: ***Impacts of Policies on Poverty: Axioms for Poverty Measurement.***

EASYPol Module 032: ***Equivalence Scales: General Aspects***.

The trainer may also consider the opportunity to present the relevant segments of the country case study based on real data in EASYPol Module 042: ***Inequality and Poverty Impacts of Selected Agricultural Policies: The Case of Paraguay.***

6.3 Frequently asked questions

- ✓ How does the Lorenz Curve vary when the original income distribution is changed?
- ✓ What is the meaning of Lorenz dominance?
- ✓ Does the Lorenz Curve take into account the size of income or only the way in which income is distributed?
- ✓ What happens if Lorenz Curves intersect?
- ✓ Against which benchmark does Lorenz Curve measure inequality?

APPENDIX - LORENZ CURVE PROPERTIES

As the Lorenz Curve is a relation between cumulative proportions, its initial point has coordinates (0,0) and its final point has coordinates (1,1). If each individual had the same income the Lorenz Curve would be equal to the equidistribution line. Since incomes are not equal and poor individuals own proportionally less income than rich people, the Lorenz Curve will lie below the equidistribution line. In fact, in a standard income distribution the Lorenz Curve is a convex curve.

Property 1

The Lorenz Curve is a convex curve. The reason may be given in an intuitive way, considering that individual incomes y_i, in the Lorenz Curve, are lined up in ascending order. In other words, in a Lorenz Curve, it must be:

$$y_1 \le y_2 \le ... \le y_n$$

Now, let us define the cumulative proportion of income and population as follows:

$$q_i = \frac{y_1 + y_2 + ... + y_i}{y_1 + y_2 + ... + y_n} = \frac{y_1 + y_2 + ... + y_i}{Y} \rightarrow \text{cumulative proportion of income}$$

$$p_i = \frac{i}{N} \rightarrow \text{cumulative proportion of population}$$

The q_i 's are therefore the values of the y-axis of the Lorenz Curve, while the p_i's are the corresponding values of the x-axis. As incomes are lined up in ascending order, it must be that each q is greater than the q immediately before, i.e.:

$$q_1 \le q_2 \le ... \le q_n$$

The way in which the q's are defined, it follows that:

$$q_i - q_{i-1} = \frac{y_1 + y_2 + ... + y_{i-1} + y_i}{Y} - \frac{y_1 + y_2 + ... + y_{i-1}}{Y} = \frac{y_i}{Y}$$

i.e., any given income share is equal to the income shares cumulated until it subtracts all income shares cumulated beforehand. It thus follows that with incomes y lined up in ascending order it must be that:

[A.1] $$q_1 - q_0 \le q_2 - q_1 \le q_3 - q_2 \le ... \le q_n - q_{n-1}$$

where q_0=0.[8]

Now, let us consider what happens in Figure A-1, below, where only the initial part of the Lorenz Curve corresponding to a hypothetical income distribution is depicted. The points q_1, q_2 and q_3 represent the cumulative heights of the Lorenz Curve for the first three individuals of this distribution. Let us now assume that instead of ending with the

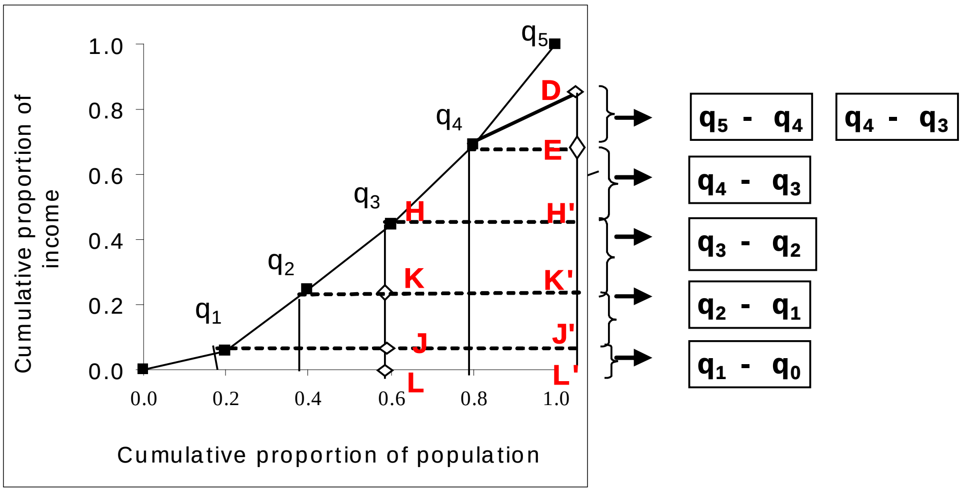
[8] This is a technical addition that does not change the result. Graphically, it allows the Lorenz curve to start from the origin.

small squared black box, the Lorenz Curve would end with the line joining q_3 and q_4, introducing a non-convexity in the Lorenz Curve.

Each height may be decomposed as the difference between it and the height of the point before it. For example, height q_3 (equivalent to HL in Figure A-1) can be decomposed in the segment **HK** (the difference between q_3 and q_2), the segment **KJ** (the difference between q_2 and q_1) and the segment **JL** (the difference between q_1 and q_0). Analogously, height q_5 can be decomposed in the segment **DE** (the difference between q_5 and q_4), **EH'** (the difference between q_4 and q_3), **H'K'** (the difference between q_3 and q_2), **K'J'** (the difference between q_2 and q_1) and **J'L'** (the difference between q_1 and q_0).

What is wrong with this latter decomposition? It does not satisfy condition [A.1]. By the way the Lorenz Curve is built, we know that condition [A.1] **must** always be satisfied. In Figure A-1 it is clearly the case that q_5-q_4 is lower (and not greater as it should be) than q_4-q_3, which contradicts condition [A.1]. This occurs because we have included non-convex segments in the Lorenz Curve. As condition [A.1] derives from the way the Lorenz Curve is built and since a non-convex segment violates condition [A.1], the Lorenz Curve cannot have non-convex segments.

Figure A-1: The Lorenz Curve is convex



By the way q_i and p_i are defined, it also clear that the initial coordinates of the Lorenz Curve (p_0, q_0) are at the point **(0,0)**, while the maximum coordinates of the Lorenz Curve (p_n, q_n) are at the point **(1,1)**.

Property 2

A useful property of the Lorenz Curve, which will be important later on, is that it is invariant to equi-proportional changes of the original distribution but not to equal absolute changes.

When all incomes are scaled by the same percentage factor, the Lorenz Curve does not change. This is so because scaling all incomes by the same percentage will also increase total incomes by the same percentage. Therefore, the heights of the Lorenz Curve (the

q's) do not change as all members at the numerator and the denominator are multiplied by the same constant:

$$q_i = \frac{y_1(1.2) + y_2(1.2) + ... + y_i(1.2)}{y_1(1.2) + y_2(1.2) + ... + y_n(1.2)} = \frac{(1.2)[y_1 + y_2 + ... + y_i]}{(1.2)Y} = \frac{[y_1 + y_2 + ... + y_i]}{Y}$$

When all incomes are added or subtracted by the same absolute amount, the Lorenz Curve does instead change, as in that case the numerator and the denominator of the heights q do not change in the same proportion.

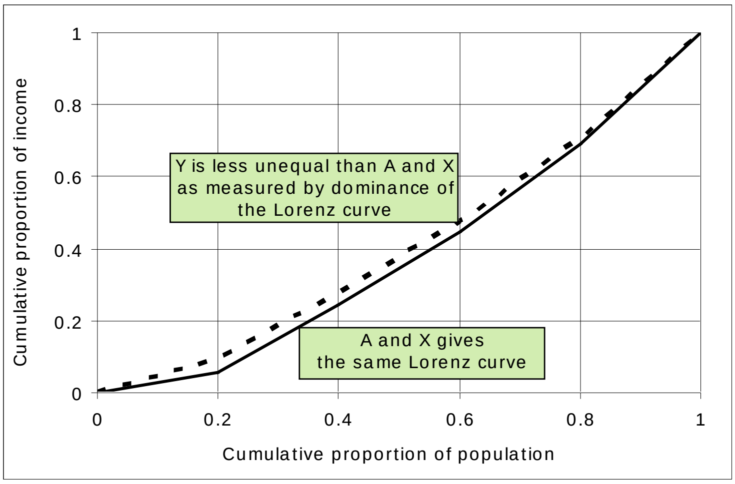
Table A-1, below, illustrates the case, considering income distribution **A** and two alternative distributions obtained from **A**, the first by multiplying all incomes in **A** by 1.2 (distribution **X**) and the second by adding 2,000 to all incomes in **A** (distribution **Y**).

Table A-1: The Lorenz Curve is scale invariant

Individuals	Income distribution A	Income distribution X	Income distribution Y
1	2,417	2,900	4,417
2	7,800	9,360	8,800
3	8,489	10,187	9,489
4	10,072	12,086	11,072
5	12,957	15,548	13,957
Total	41,735	50,082	47,735

Remark: Lorenz curves takes into account how income is distributed. **They do no take into account** if income distributions have different total incomes.

See Module on Generalised Lorenz Curves[10].



The graph associated to Table A-1 shows the outcome in terms of theLorenz Curve. Two of them perfectly superimpose (**A** and **X**), while the other (**Y**) is closer to the equidistribution line than the other two. This property is formally defined by saying that the **Lorenz Curve is scale invariant and not translation invariant.**[9] Whenever

[9] See EASYPol Module 008: : ***Impacts of Policies on Poverty: Axioms for Poverty Measurement***.

[10] EASYPol Module 002: ***Social Welfare Analysis of Income Distributions: Ranking Income Distributions with Generalised Lorenz Curves.***

incomes are increased by adding an equal amount of money, the new Lorenz Curve would be closer to the equidistribution line than the original one. Adding equal amounts would therefore generate less inequality in Lorenz Curve terms. The opposite holds true if income were decreased by the same amount. Subtracting equal amounts would therefore generate more inequality in Lorenz Curve terms.

7 REFERENCES AND FURTHER READINGS

Lorenz M.O., 1905. Methods of Measuring the Concentration of Wealth, *Journal of the American Statistical Association* (new series) **70**, pp. 209-217.

Further readings, with different levels of technicalities, are:

Cowell F.A., 1977. *Measuring Inequality*, Philip Allan, Oxford, UK.

Lambert P., 1993. *The Distribution and Redistribution of Income*, Manchester University Press, Manchester, UK. 2nd edition (especially Chapter 2 – The Size Distribution of Income);

Sen A., 1997. *On Economic Inequality*, Clarendon Press, Oxford, UK (especially Chapter 2 – Measures of inequality).