# INEQUALITY AND ITS MEASUREMENT

Satya R. Chakravarty

Indian Statistical Institute,
Kolkata, INDIA

Home Page: www.isical.ac.in/~satya

# Introduction

The purpose of the lecture of today is to examine some selected aspects of distributions of income across populations.

We will analyse the concept of inequality of income and its different measures developed in the literature.

Many of you have been attending the “Inequality and…?” lectures since 2013.

Have you ever asked:

**WHY INEQUALITY and HOW SPECIALISTS MESURE IT?**

An index of income inequality is a scalar measure of interpersonal income differences within a given population.

High income inequality means concentration of high incomes in the hands of few.

Economic growth may be affected by the inability of many to invest in education and their lower health levels, among other factors.

Large wealth gaps can give rise social conflicts, and higher security costs, for both businesses and governments.

In terms of social outcomes, inequality has impacts on several issues, including, health, education, incidence of crime and violence (Deaton, 2001).

# THE ROLE OF MIDDLE CLASS

High income inequality is likely to compress the size of the middle class.

A large and rich middle class contributes significantly to the well-being of a society in many ways, particularly, in terms of high economic growth, better health status, higher education level, a sizeable contribution to the country's tax revenue and a better infrastructure, and more social cohesion resulting from fellow feeling.

In the words of Aristotle (-350)

'…the best political economy is formed by citizens of the middle class, and that those states are likely to be well-administered, in which the middle class is large.'

On the other hand, a society characterized with a small middle class and more persons away from the middle income group may lead to a strained relationship between the subgroups on the two sides of the middle class which can generate unrest.

# INEQUALITY MEASURES

Inequality indices are employed to address a wide range of issues in development studies.

Some of the standard questions that arise in this context are:

- Is the income distribution in the country more equal now than it was five years ago?
- Is inequality in region A of the country is more/less than that in region B?
- How much inequality is due to differences between racial/education/gender groups? And how much is due to differences between each group?

# Measuring Inequality

Since often we will compare inequality levels of distributions with different population sizes, we will restrict our attention on the set of all possible income distributions, which we denote by $D$.

For instance, we may need to compare the inequality of the distribution $x=(1,3)$, two individuals, one with income 1 and the second with income 3, with that of the distribution $x=(2,4,6)$, with three individuals with incomes 2, 4, 6. Both these profiles are in the set $D$.

Thus, if we consider the Indian states in which population sizes are different, then $D$ will contain all the income distributions in the different states in India.

We assume that the income distributions are non-decreasingly ordered (first smallest income, last biggest income). We denote the mean of a distribution by $\lambda$ . For example for x=(1,3), $\lambda$ is 2.

By an inequality index $I$ we mean a real valued function defined on $D$. That is, for any distribution $x$ in $D$, $I(x)$ is a real number indicating the level of inequality of $x$.

Not every function is an appropriate measure of inequality. For example a constant function would be useless since it would give the same number independently of the distribution.

The way social scientists proceed is by assuming that $I$ satisfies some properties, also known as axioms.

We generally use **four axioms** for a good measure of inequality that allows us to compare inequality for example:

between Luxembourg and China (very different population size!)

and

between Luxembourg and Mozambique (very different total income!).

## *Axiom 1: The population replication invariance principle*

To consider the first postulate, suppose it **is** necessary to decide which distribution is more unequal between:

$$x^{\bullet} = (1,3) \quad \text{and} \quad x^{0} = (2,4,6)$$

Note that these distributions have different population sizes, 2 individuals in the first and 3 individuals in the second.

A meaningful comparison of inequality across distributions requires that the underlying population sizes should be the same.

If we view inequality as an average concept, that is based on population shares independently on the actual number of individuals, then an income by income replication of the distribution should keep inequality unchanged.

Hence the way social scientists proceed in the comparison of populations with different number of individuals is to assume that replicating the population does not alter inequality.

That is, inequality in the two cases below is the same:

$$x^{\bullet} = (1,3) \qquad \hat{x} = (1,1,1,3,3,3)$$

Following this property, we replicate the two original distributions enough times to arrive to two final distributions with the same number of individual:

From $x^{\bullet} = (1,3)$ $x^{0} = (2,4,6)$

To $\hat{x} = (1,1,1,3,3,3)$ $\tilde{x} = (2,2,4,4,6,6)$

The final distributions have a common population size 6 and we compare these two.

If the replicated populations are equally unequal respectively to their original counterparts, then we say that the inequality index is population invariant.

That is, replicating the population does not alter inequality.

Formally, $I$ is population replication invariant if $I(x)=I(y)$, where $y$ is any $k$-fold replication of $x$, $k$ being any positive integer greater than 1.

Thus, the comparison of inequality of two original distributions with different population sizes is same as that of their replicated forms with a common population size.

Note the *mean income* is also population replication invariant.

*Axiom 2: Scale invariance*

Next, we note that the sizes of the distributions $x^{\bullet}$ and $x^{0}$ are different, that is, their total incomes are not the same.

$x^{\bullet} = (1,3)$ total is 4 $\qquad x^{0} = (2,4,6)$ total is 12

The way social scientists proceed in this situation is to assume that inequality does not change if all incomes are multiplied or divided by the same number.

That is, social scientists multiply or divide all the incomes of one distribution to obtain a transformed distribution with the same total income of the other.

Start from:

$$\hat{x} = (1,1,1,3,3,3) \quad \text{and} \quad \tilde{x} = (2,2,4,4,6,6)$$

Multiply all incomes of the first by $\frac{\lambda(x^0)}{\lambda(x^*)} = \frac{4}{2}$, to obtain $\breve{x} = (2,2,2,6,6,6)$

that has now the same total income of $\tilde{x} = (2,2,4,4,6,6)$

Social scientists will compare

$$\breve{x}=(2,2,2,6,6,6) \quad \text{and} \quad \tilde{x}=(2,2,4,4,6,6)$$

that now have the same number of individuals and the same total income.

This last property of invariance with respect to multiplication/division is known as scale invariance.

Formally, an inequality index $I$ possesses the *scale invariance property*, if $I(x)= I(y)$ , where $y$ is obtained by multiplying $x$ by a positive scalar.

Thus, under the scale invariance and population replication invariance postulates, the inequality comparison of the distributions with different population sizes and different totals is equivalent to that of the transformed distributions having the same total income and the same population size.

*Axiom 3: Symmetry or Anonymity*

It is also desirable that in the measurement of income inequality the individuals should not be distinguished by anything other than their incomes.

This condition is satisfied if the inequality index satisfies *symmetry*, which demands that inequality should be insensitive to reordering of the incomes.

That is, inequality in (1,2,3) is the same as inequality in (2,1,3), in (3,1,2), etc..

Symmetry allows us to define the inequality index directly on ordered income distributions (as we have done), that is, on vectors where the smallest income is in the first position and the largest in the last.

*Axiom 4: The principle of transfers*

Let us consider the two distributions

$x^A = (2,4,5,9)$ and $x^B = (2,4,6,8)$ of the same total income, 20.

In both distributions incomes are unequally distributed. But the richest person enjoys a higher income in $x^A$ than in $x^B$ and the opposite is true for the third richest person. In fact, $x^A$ is obtained from $x^B$ by a transfer of 1 unit of income from the richest to the third richest.

The difference between the two distributions is one progressive transfer.

In general, a progressive transfer of income, that is, a transfer of income from a person to anyone who has a lower income so that the donor does not become poorer than the recipient, should reduce inequality.

This property is known as *the principle of transfers* (the Robin Hood principle).

The principle of transfers is the key axiom of inequality measurement.

Inequality generates spread in the distribution and a progressive transfer reduces the spread.

$x^A = (2,4,5,9)$ $\qquad\qquad$ $x^B = (2,4,6,8)$

$x^A$ is more unequal than $x^B$

The indices of inequality social scientists like to use satisfy these four axioms:

1) The population principle
2) Scale invariance
3) Anonymity or Symmetry
4) The principle of transfers

Several indices do.

# Inequality Indices

The most widely used index of inequality that obeys the principle of transfers, symmetry, the population replication invariance principle and the scale invariance condition is the **Gini index**, which is defined for any income distribution $x = (x_1, x_2, .., x_n)$ as

# The Gini index

$$I_G(x) = \frac{1}{2n^2\lambda(x)} \sum_{i=1}^{n} \sum_{j=1}^{n} \left| x_i - x_j \right|$$

Since the incomes are non-decreasingly ordered, we can rewrite it as

$$I_G(x) = 1 - \frac{1}{n^2\lambda(x)} \sum_{i=1}^{n} \left(2(n-i)+1\right) x_i$$

# The coefficient of variation

Two other well-known inequality indices that satisfy these postulates are **coefficient of variation** $I_{CV}(x)$ and the **Atkinson (1970) index** $I_\theta(x)$ . For the latter index, all incomes are positive. For the other two zero, even negative incomes, are allowed(mean positive).

The coefficient of variation:

$$I_{CV}(x) = \frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(x_i - \lambda(x)\right)^2}}{\lambda(x)}$$

# The Atkinson index

$$I_\theta(x) = \begin{cases} 1 - \dfrac{\left(\dfrac{1}{n}\displaystyle\sum_{i=1}^{n} x_i^\theta\right)^{\frac{1}{\theta}}}{\lambda(x)}, \ \theta < 1,\ \theta \neq 0, \\ 1 - \dfrac{\displaystyle\prod_{i=1}^{n} (x_i)^{\frac{1}{n}}}{\lambda(x)}, \ \theta = 0. \end{cases}$$

The constant $\theta$ in the above formula is a transfer sensitivity parameter in the sense that a progressive transfer will reduce inequality by a larger amount the lower the income of the recipient of the transfer. For a given income distribution $x$ , an increase in the value of the parameter decreases inequality more.

Now, different inequality indices may not rank two distributions of income in the same way.

For instance, according to the Gini index and the coefficient of variation, the distributions $x^E = (1,5,6,8)$ and $x^A = (2,4,5,9)$ are equally unequal, whereas for any finite $\theta < 1$, $I_\theta$ regards $x^E$ as more unequal than $x^A$.

Tool that enables us to check if distributions are ranked in the same way!

# The Lorenz ordering

An ordering based on the well-known Lorenz curve can be used for checking whether different inequality indices can rank alternative distributions of income in the same way.

The Lorenz curve plots cumulative shares of total income against cumulative population shares.

# The Lorenz ordering

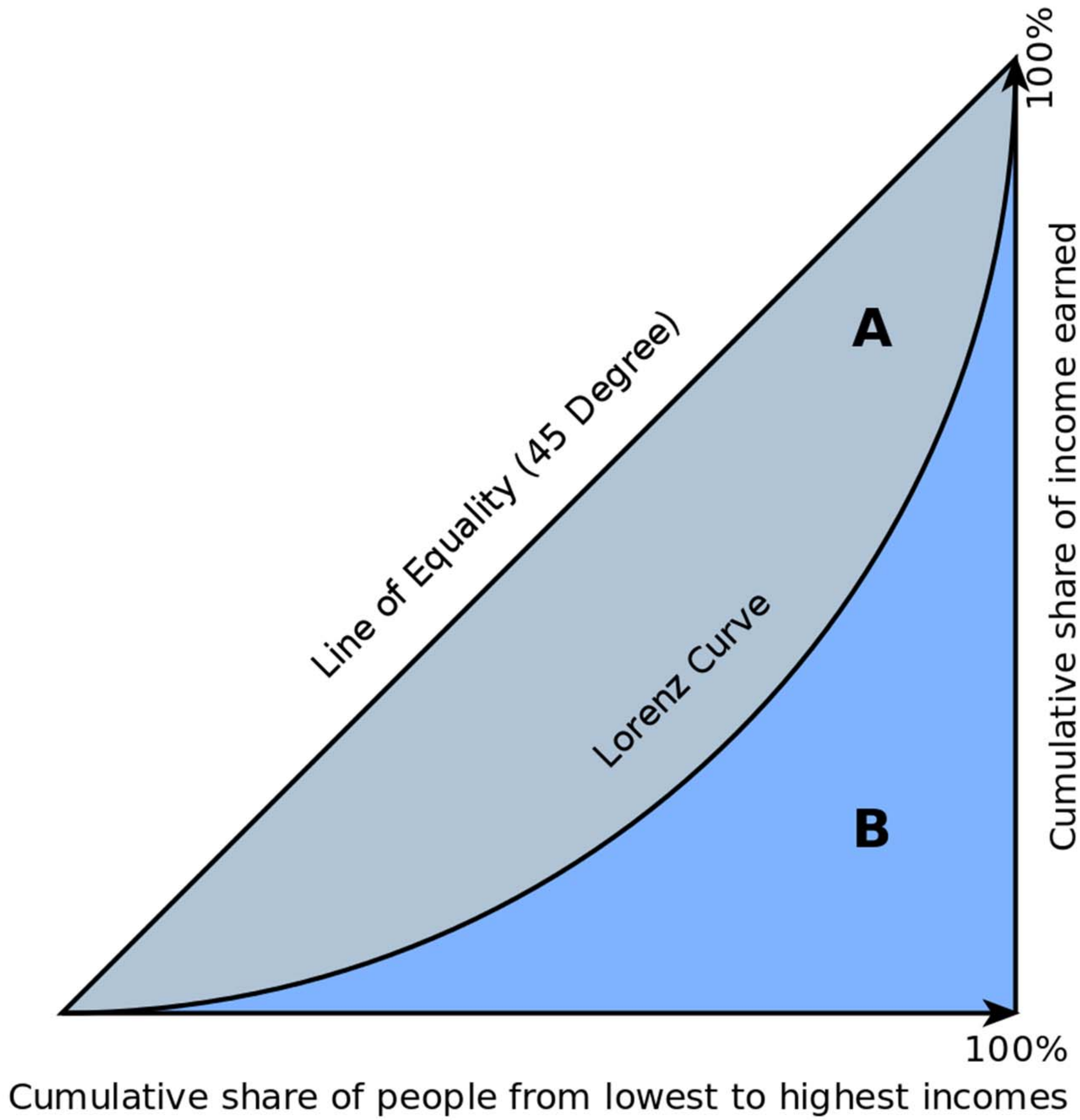
Given that 0% of the population enjoys 0% of the total income and 100% of the population possesses the entire income, the curve starts from the south-west corner with coordinates (0,0) of the of unit square and terminates at the diametrically opposite north-east corner with coordinates (1,1).

In the case of perfect equality, every $p$% of the population enjoys $p$% of the total income and the curve coincides with the diagonal line of perfect equality.

In all other cases the curve will lie below the egalitarian line.

If there is complete inequality, which is characterized by the situation where only one person has positive income and all other persons have 0 income, the curve will run through the horizontal axis until we reach the richest person and then it rises perpendicularly.

The Lorenz curve is quite useful because it shows graphically how the actual distribution of incomes differs from the egalitarian situation.


100%
Cumulative share of income earned
Line of Equality (45 Degree)
A
Lorenz Curve
B
100%
Cumulative share of people from lowest to highest incomes

In particular, the Lorenz curve allows us to rank distributions according to inequality and say that the country with Lorenz curve closer to the diagonal has less inequality than the country with Lorenz curve further away.

Formally, for any two non-decreasingly ordered distributions $x$ and $y$ , $x$ Lorenz dominates $y$ if the Lorenz curve of $x$ lies nowhere below that of $y$ and at some places (at least ) lies above.

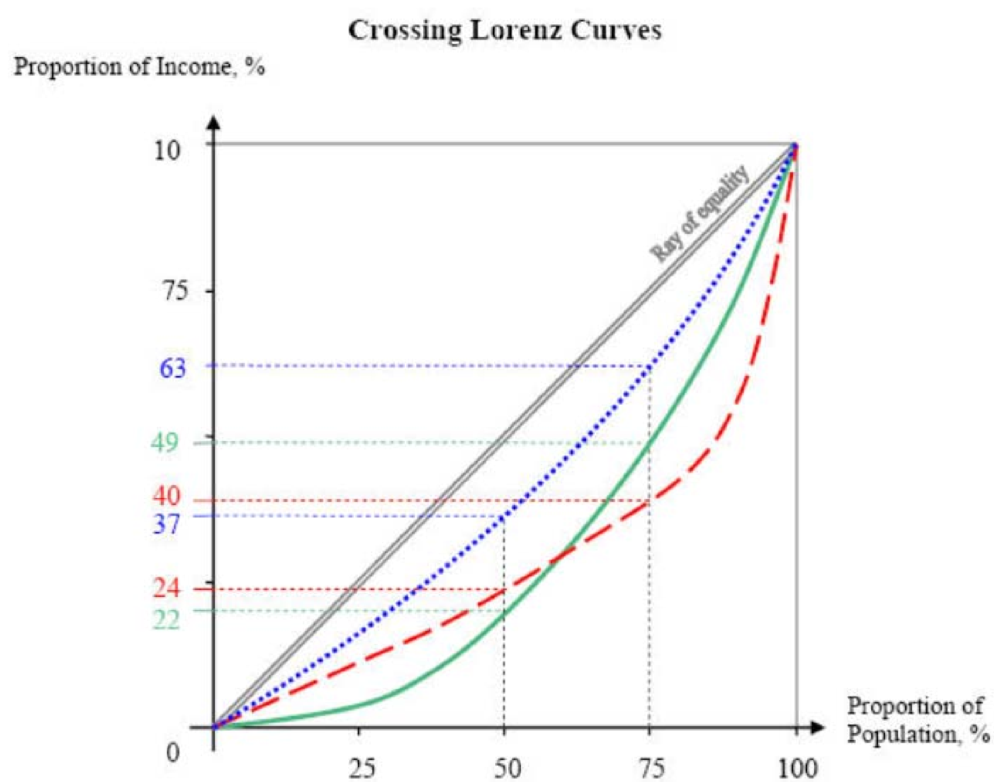
The literature established the following result:

**Theorem 1:** For any two income distributions $x$ and $y$ , $x$ Lorenz dominates $y$ if and only if $I(x) < I(y)$ for all regular inequality indices $I$, that is, those that satisfy the four axioms we discussed.

Thus, once the Lorenz domination condition holds, we can say that the former distribution is less unequal than the latter by all regular inequality indices.

Calculation of indices for the purpose of ranking is not required.

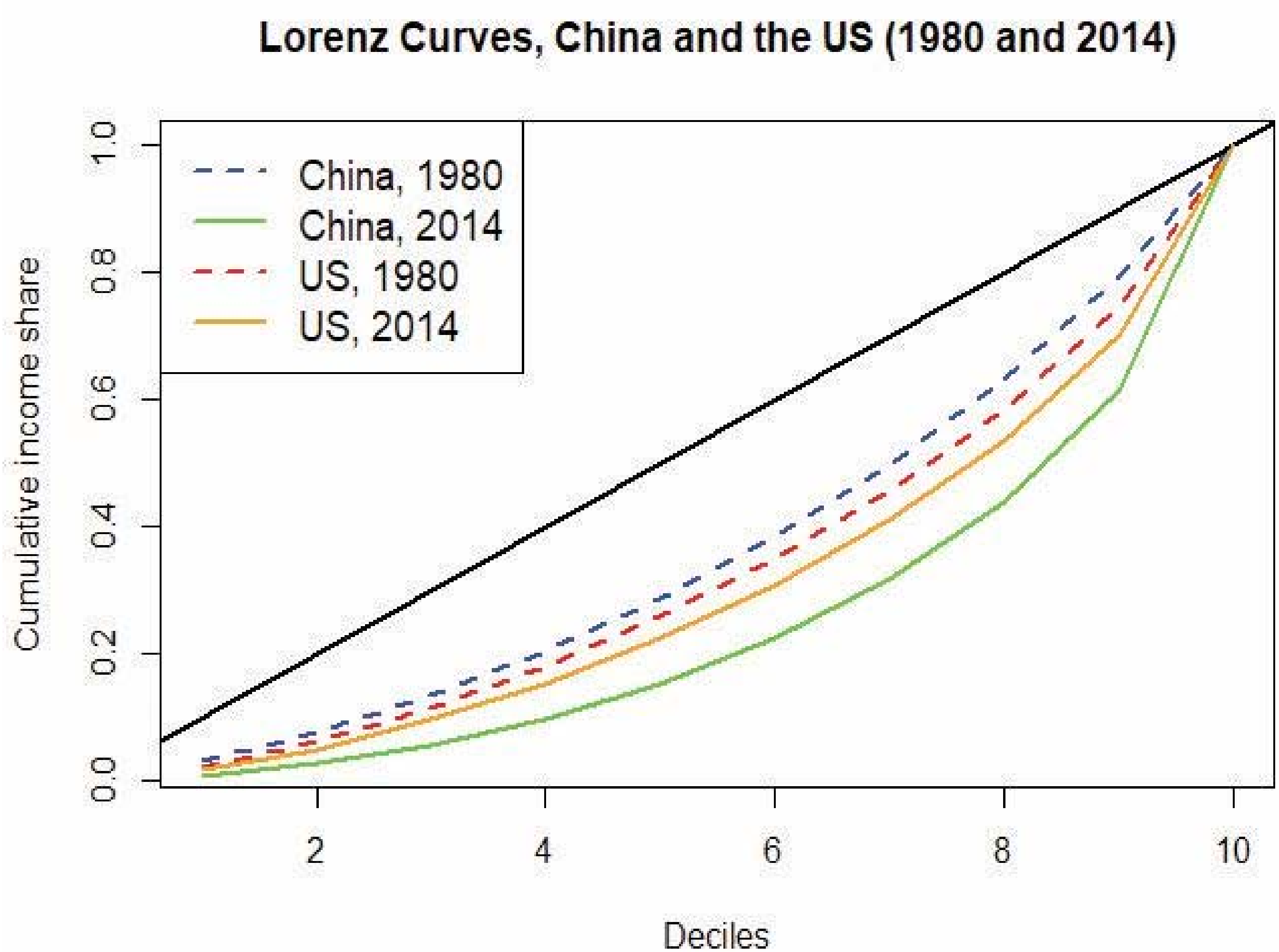
However, if the Lorenz curves of the two distributions cross, then such an unambiguous conclusion about inequality ordering cannot be drawn. There may exist two different inequality indices that will rank the distributions differently.



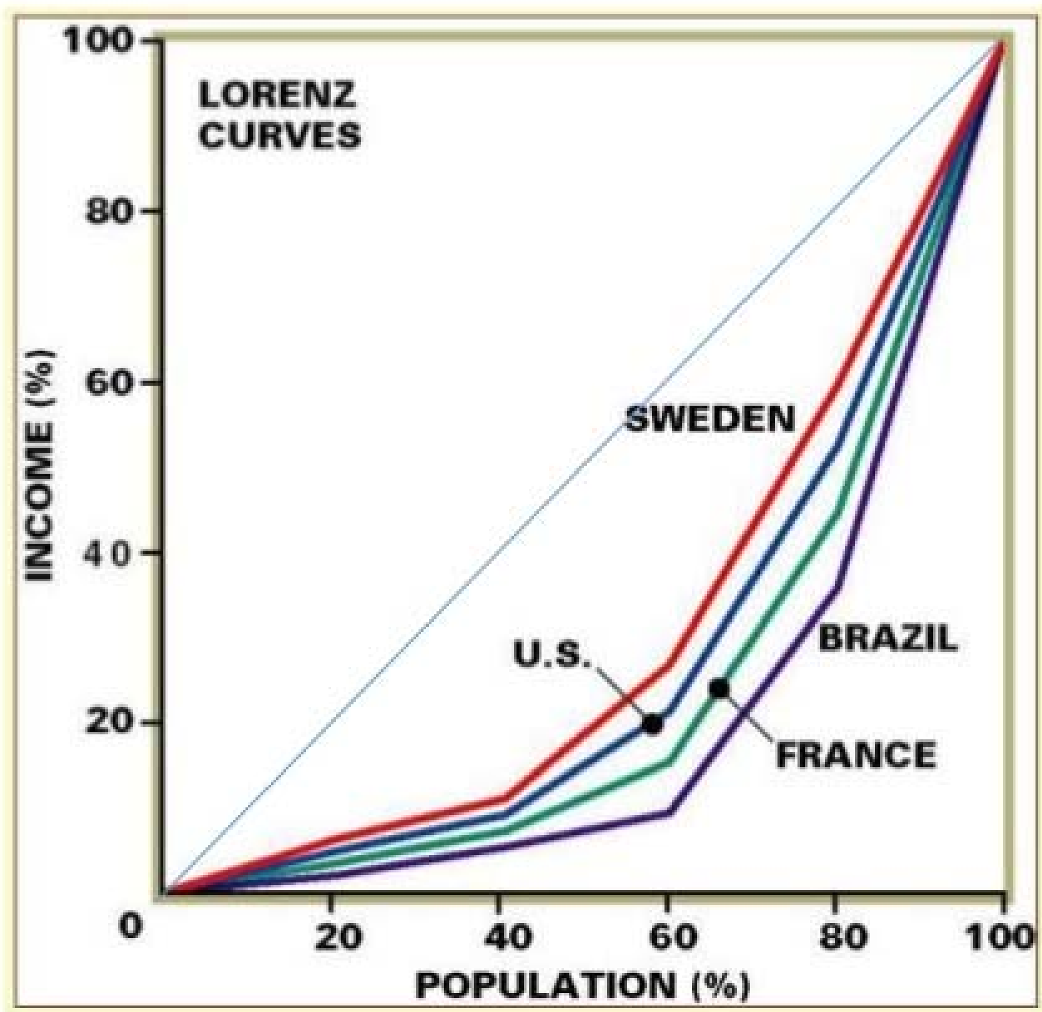
# Measuring Inequality.

For instance, the Lorenz curves of $x^A = (2,4,5,9)$ and $x^E = (1,5,6,8)$ cross and that is why their different directional rankings by the Gini index, the coefficient of variation and $I_\theta$ arise.

# Some examples of Lorenz dominance

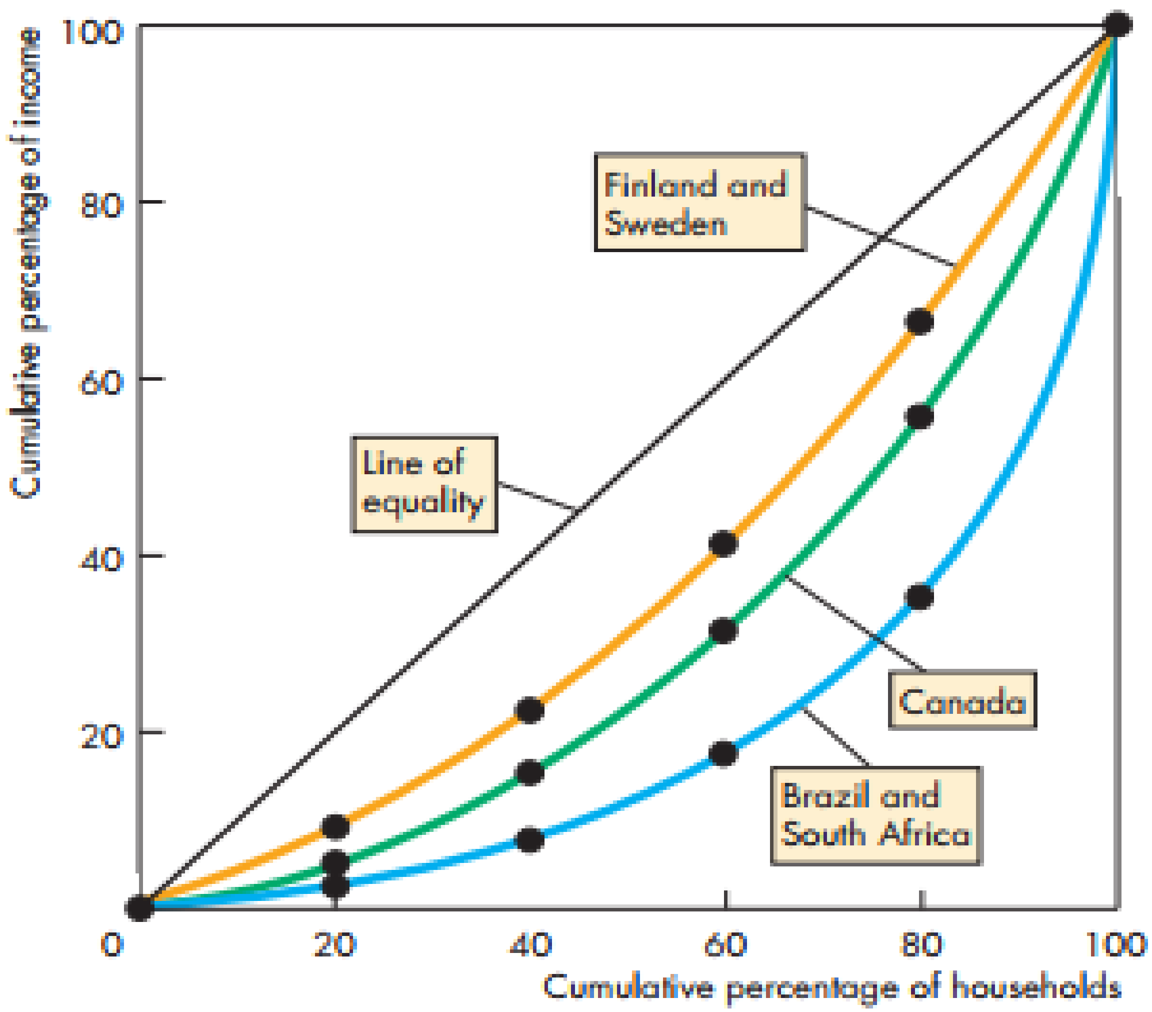


# Some examples of Lorenz dominance



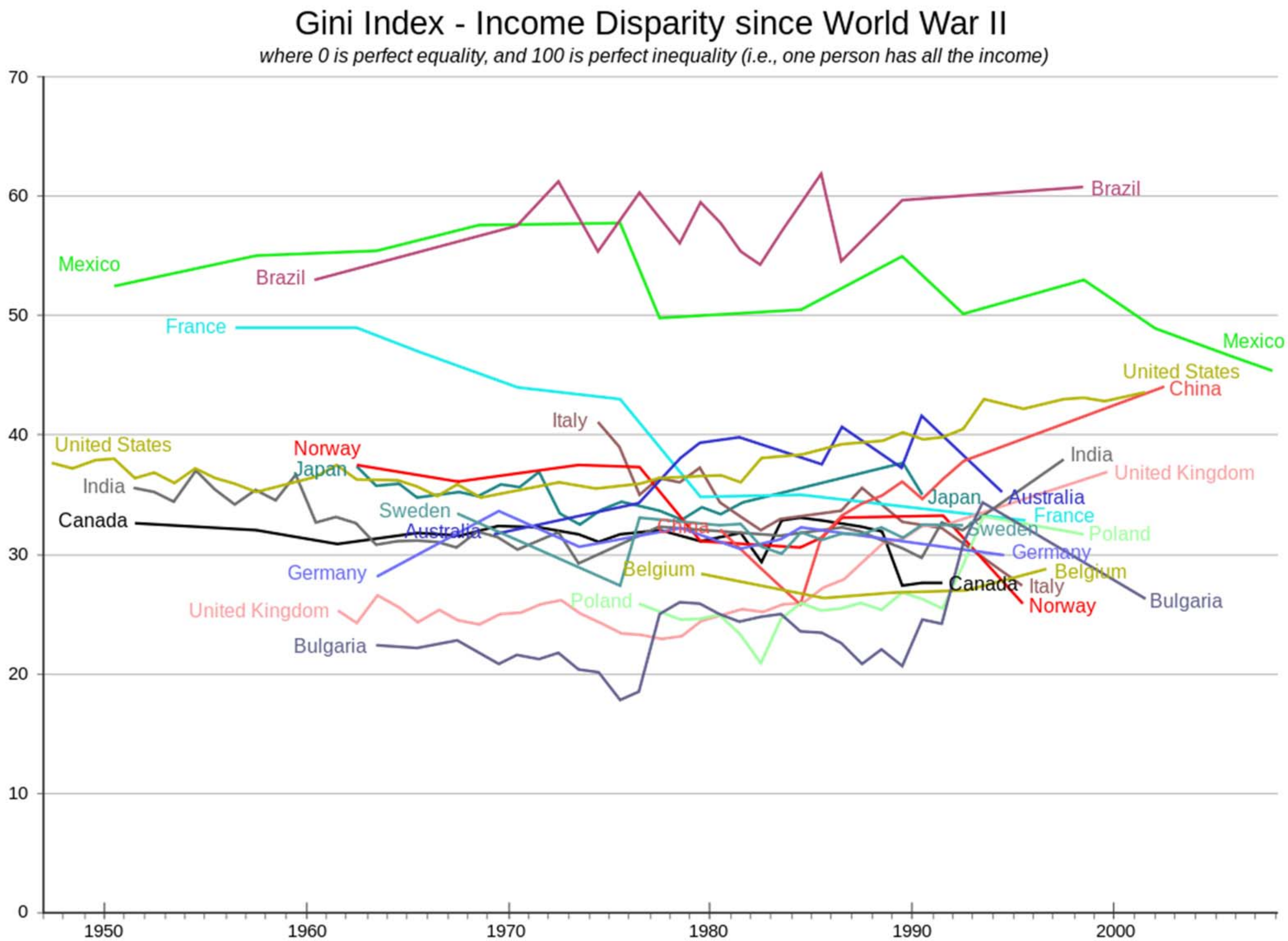
# Some examples of Lorenz dominance

FIGURE 19.7 Lorenz Curves Compared



# Examples of Gini indices

the most famous inequality index



# Caveat

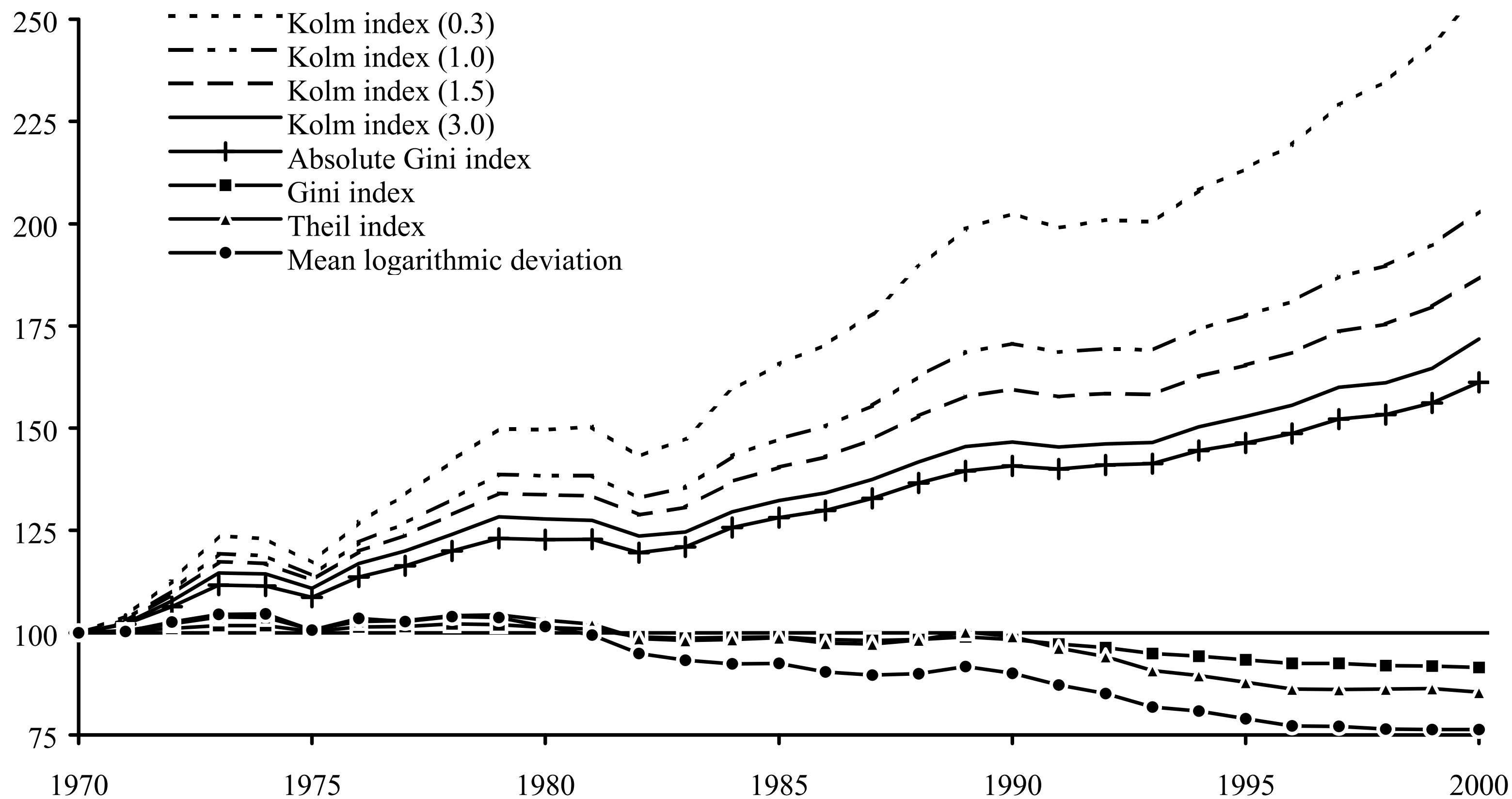
If you measure inequality using indices that satisfy other properties, the results may change.

For example, if you abandon scale invariance (invariance to multiplication/division) for translation invariance (invariance to addition/subtraction) you may see a different path:

# Caveat

**INTERNATIONAL INCOME INEQUALITY, 1970-2000: RELATIVE AND ABSOLUTE INDICES**

(Indices: 1970=100)



# Final message

Inequality is complicated…