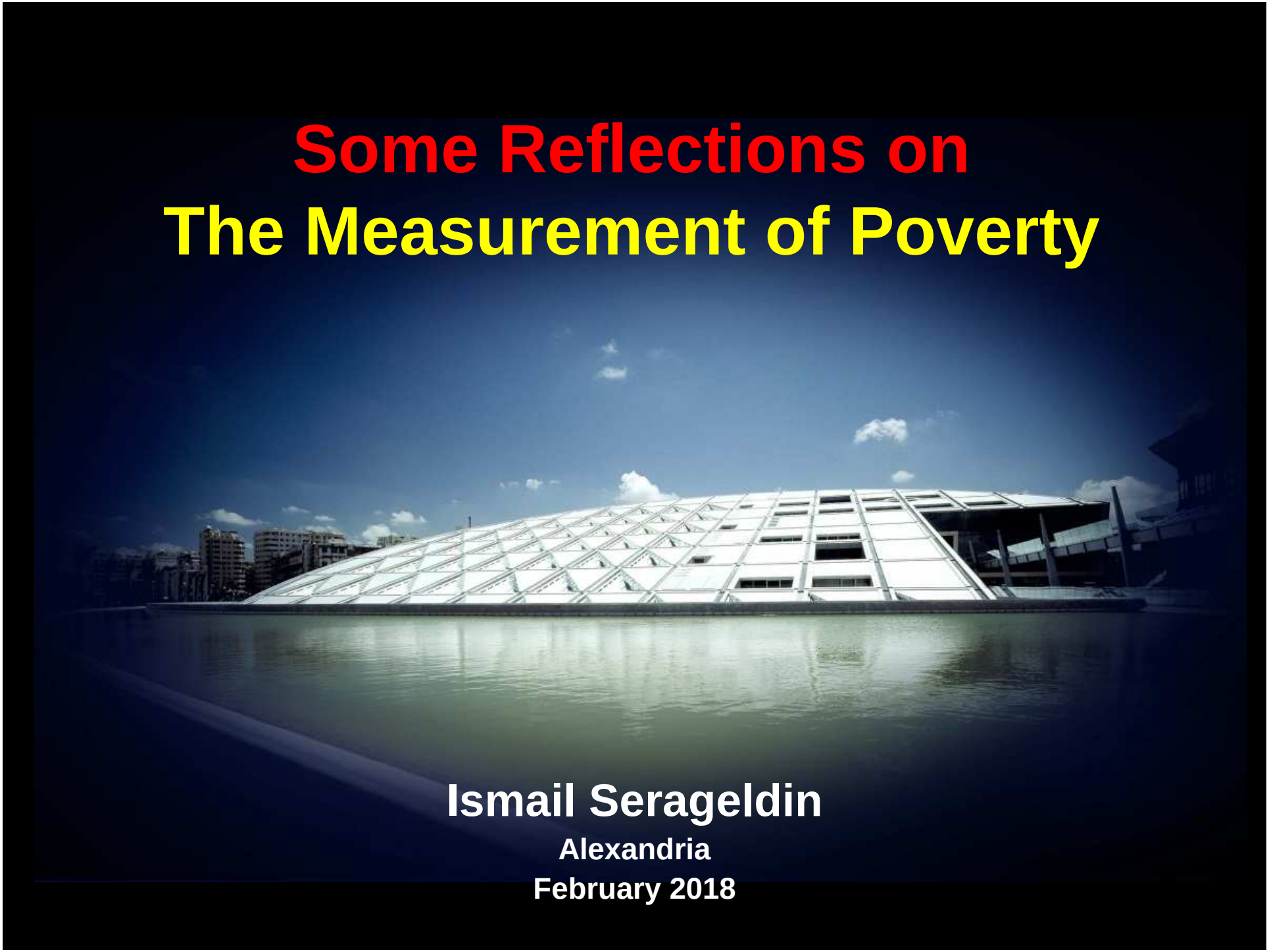
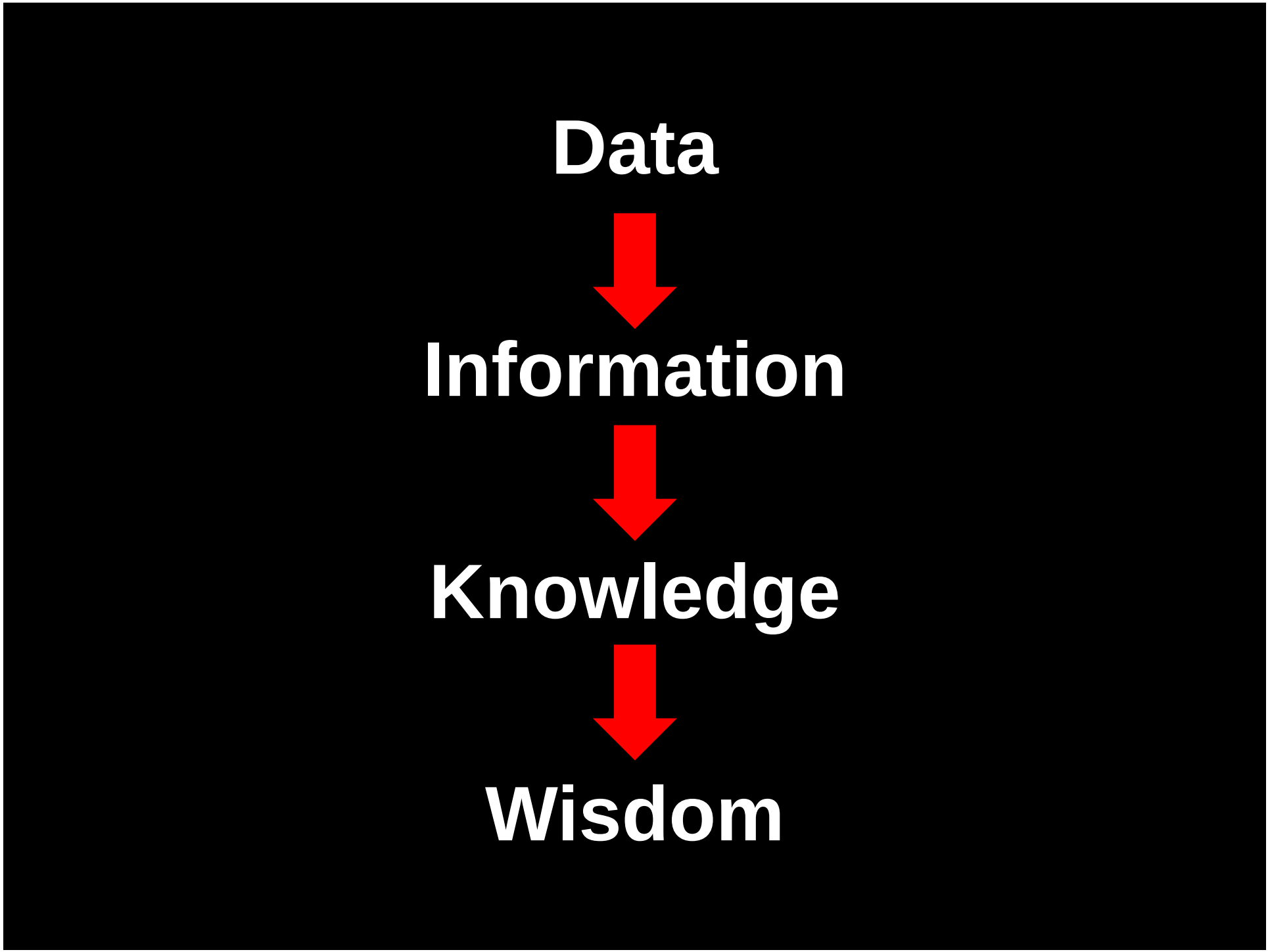
Some Reflections on
The Measurement of Poverty



Ismail Serageldin
Alexandria
February 2018

Outline

- Diagnosing Poverty
- Measuring Income
- Measuring Poverty
- Measuring Inequality
- Extreme Poverty and Resilience
- From Data to Models
- Building Models: Design and Errors
- Envoi


Data
Information
Knowledge
Wisdom

Diagnosing Poverty

Diagnosing Poverty

- On Absolute And Relative Poverty
- Deprivation, Dispossession, And Societal Marginalization
- Rural And Urban Poverty
- Problems Of The Ultra Poor
- Social Versus Economic Policies, Programs And Projects

Diagnosing Poverty

- On Absolute And Relative Poverty
- Deprivation, Dispossession, And Societal Marginalization
- Rural And Urban Poverty
- Problems Of The Ultra Poor
- Social Versus Economic Policies, Programs And Projects

Absolute and Relative Poverty

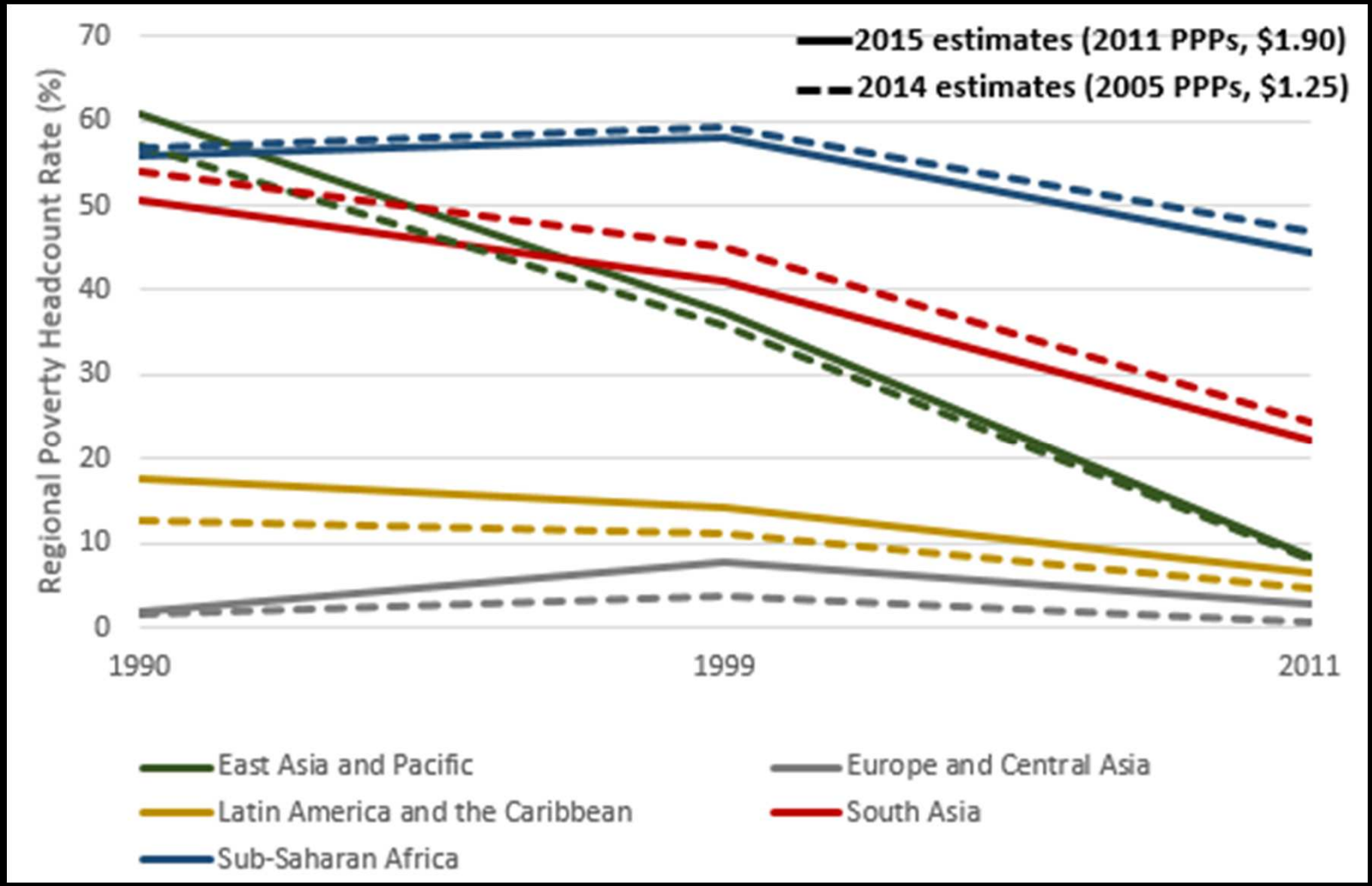
Absolute Poverty

- **Absolute Poverty** is defined as having income less than the minimum amount needed by a person or household to obtain the basic necessities for living.
- **$1/day per person** was the benchmark for international comparisons as of 1990.
- It became $1.25 / day per person in 2005.
- Now it is **$1.90 /day per person**

Calculating Poverty Lines Using PPP Exchange Rates

- Remember that PPP exchange rates are calculated so as to offset differences in absolute price levels: one PPP dollar should buy the same basket of goods in Kenya, India or the US. If prices in poor countries are lower, their currencies are stronger in purchasing power terms: the Kenyan shilling or the Indian rupee buy more (in their respective countries) than we used to think, relative to what one dollar buys in the United States.
- In other words, the US dollar's purchasing power (in the US) in 2011 PPPs is lower relative to the purchasing power of the currencies of most poor countries (in those countries). The new PPPs effectively reflect a weaker dollar, relative to the currencies of most poor countries.
- This is why a poverty line that is constant in real terms in poor countries, is now higher in US dollars. $1.90 in 2011 buys approximately the same things as $1.25 did in 2005 in poor countries, which is why poverty has changed very little. That the value is higher in US dollar terms is merely a reflection of a 'weaker" dollar in PPP terms

Percent of Population below Regional Poverty Lines



Relative Poverty

- Varies from society to society
- Sometimes taken as the lowest 40% of the income distribution in that country
- Sometimes defined as someone receiving below 60% of the median income

Almost all the hungry are among those in absolute poverty

But poverty is not just about income or money… it is more…

Diagnosing Poverty

- On Absolute And Relative Poverty
- **Deprivation, Dispossession, And Societal Marginalization**
- Rural And Urban Poverty
- Problems Of The Ultra Poor
- Social Versus Economic Policies, Programs And Projects

It is not just the absence of income that defines poverty

It is marginalization, deprivation and social exclusion

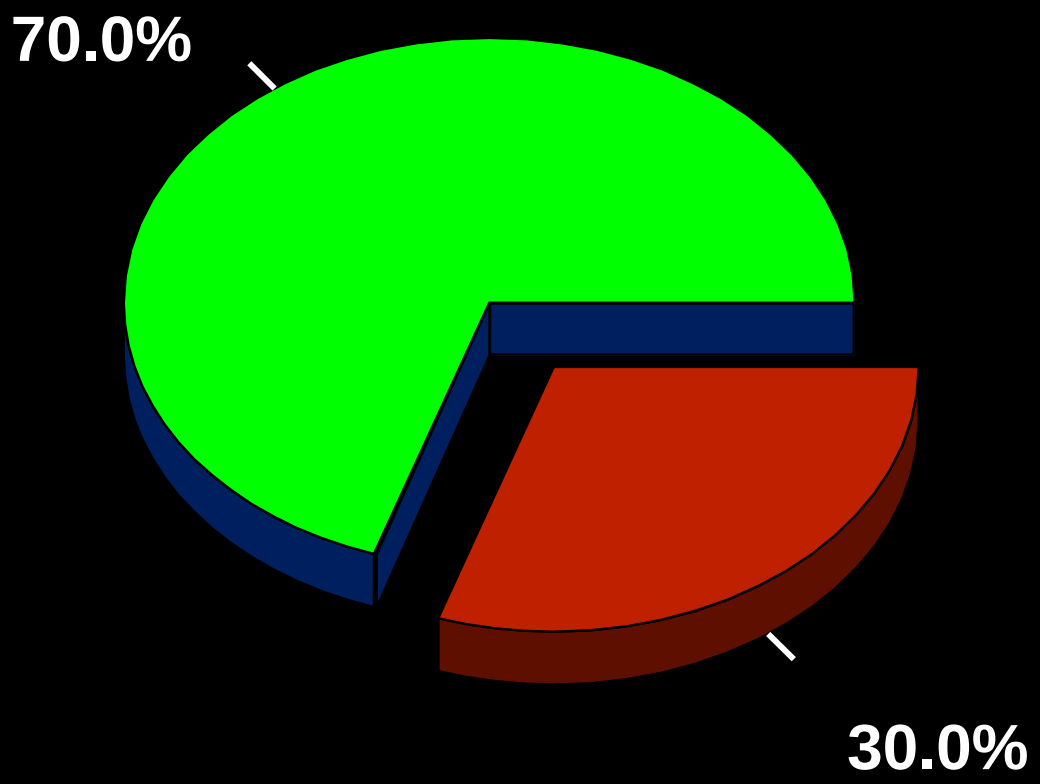
Loss of dignity

Social Exclusion

Diagnosing Poverty

- On Absolute And Relative Poverty
- Deprivation, Dispossession, And Societal Marginalization
- **Rural And Urban Poverty**
- Problems Of The Ultra Poor
- Social Versus Economic Policies, Programs And Projects

Rural and Urban Poverty in Developing Countries

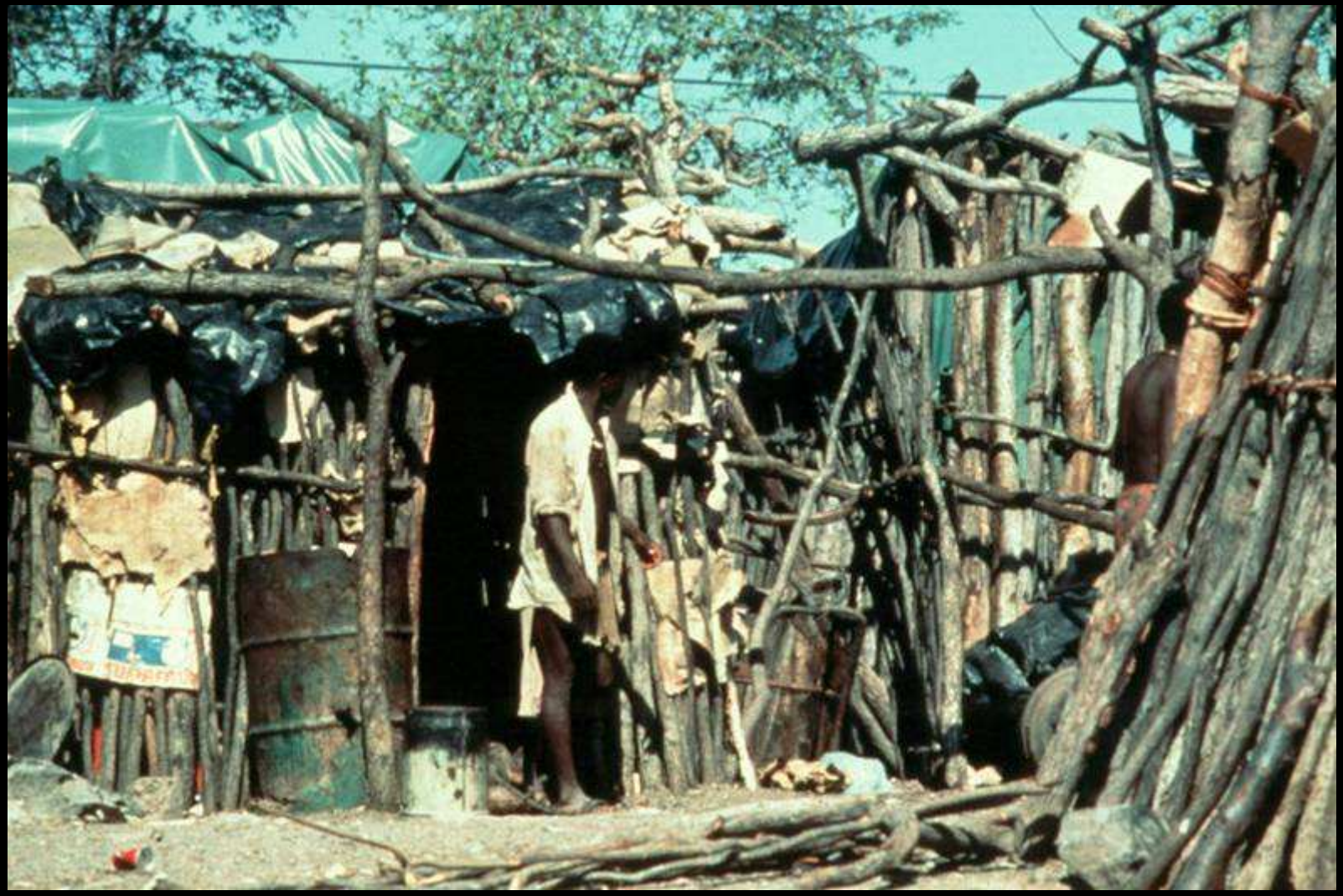


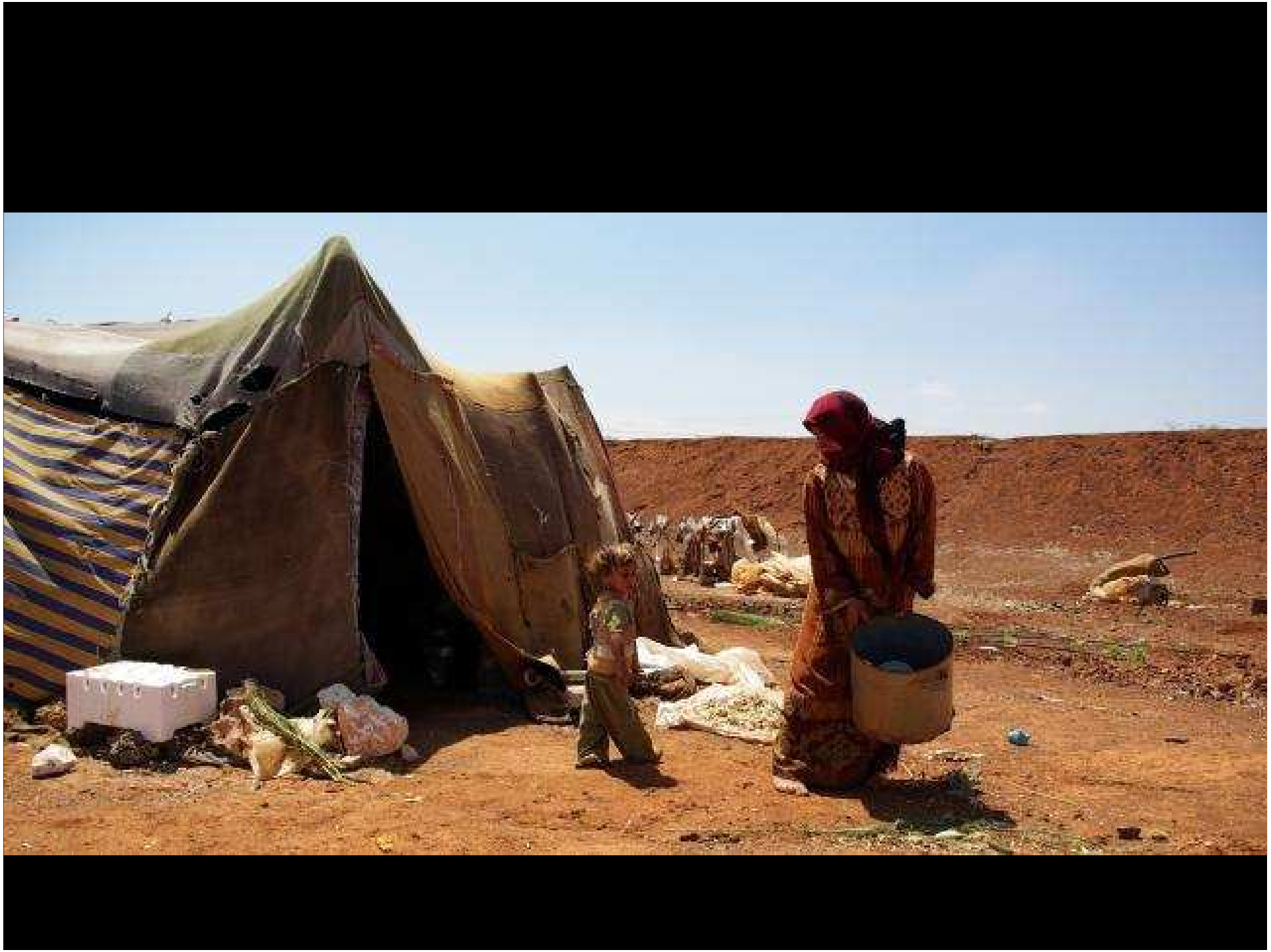
Source: IFPRI estimate from World Bank data.

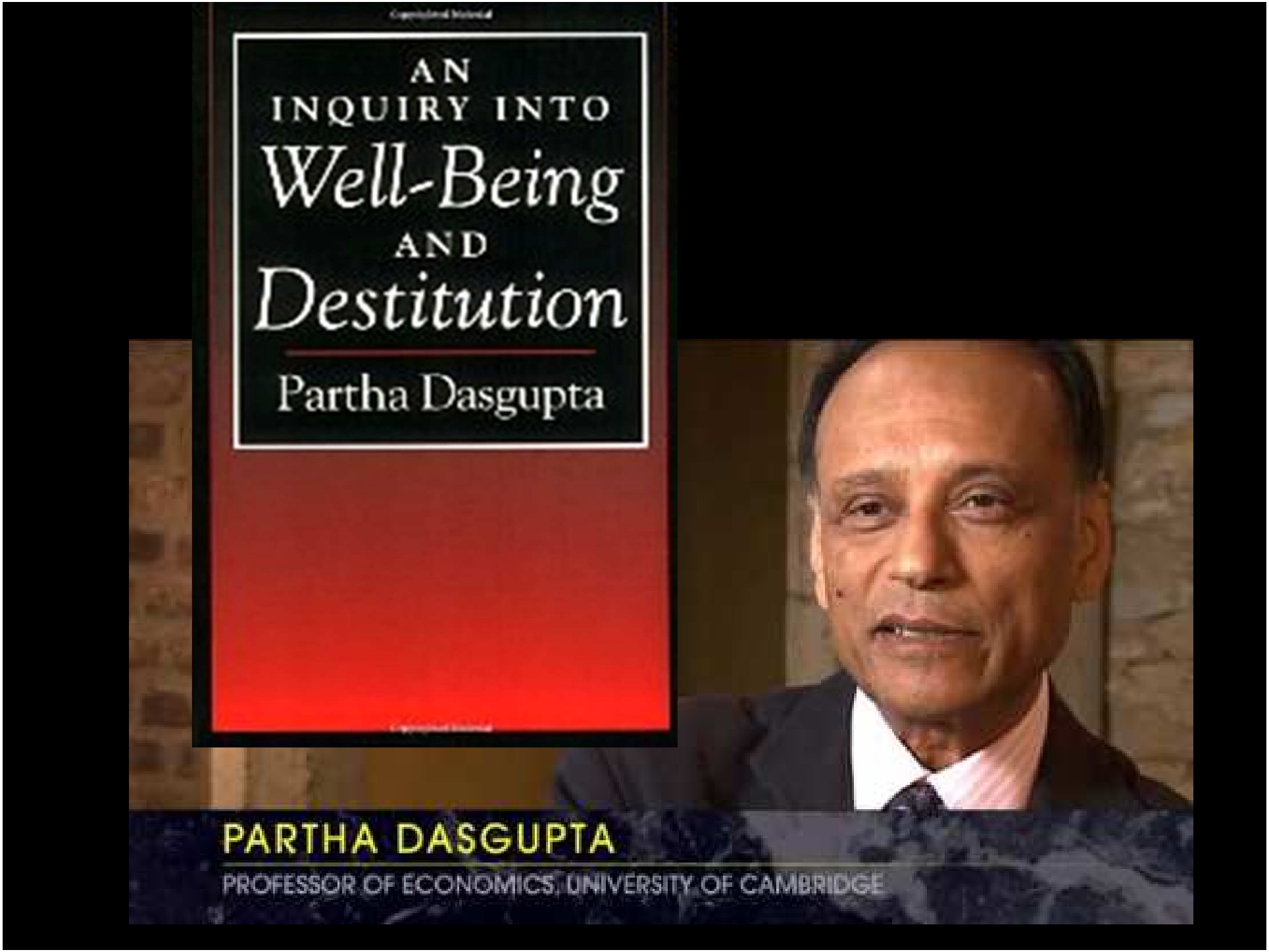
Diagnosing Poverty

- On Absolute And Relative Poverty
- Deprivation, Dispossession, And Societal Marginalization
- Rural And Urban Poverty
- **Problems Of The Ultra Poor**
- Social Versus Economic Policies, Programs And Projects

The Ultra Poor require special help






AN
INQUIRY INTO
Well-Being
AND
Destitution
Partha Dasgupta
PARTHA DASGUPTA
PROFESSOR OF ECONOMICS, UNIVERSITY OF CAMBRIDGE

Diagnosing Poverty

- On Absolute And Relative Poverty
- Deprivation, Dispossession, And Societal Marginalization
- Rural And Urban Poverty
- Problems Of The Ultra Poor
- Social Versus Economic Policies, Programs And Projects


THIS
SUCKS!

The Need for Social Inputs Into Development Decisions

- Social policy is more than the social consequences of economic policies
- Social goals and policies complement economic ones
- Economic Analysis by itself is insufficient: Social, cultural, political and ethical dimensions must be introduced

Elements Of A Social Policy - I

- To maintain social cohesion
- To foster equity
- To reach the ultra poor and other marginalized groups
- To uphold cultural identity (shared universal values and solidarity, not divisive micro-identities)

Elements Of A Social Policy - II

- To promote participation (voice, choice and empowerment through access to knowledge and resources)
- To facilitate social mobility (inter-generational, geographic and occupational)
- To support institutional development
- To enable participatory social research

Diagnosing Poverty

- On Absolute And Relative Poverty
- Deprivation, Dispossession, And Societal Marginalization
- Rural And Urban Poverty
- Problems Of The Ultra Poor
- Social Versus Economic Policies, Programs And Projects

Measuring Income

There are many kinds of poverty


†
I AM
HOMELESS
PLEASE HELP

Living in Resource-poor Environments

But in the final analysis,
we almost always go back to
income
to measure poverty

Defining Poverty

- Although we all recognize the multi-dimensional character of poverty, we almost always go back to defining poverty in terms of income
- We have much improved by using Household Surveys; but
- Despite our reservations on income as GDP/Capita, it is still widely used.

The most common measure of income is GNP / Capita

But GNP is a measure of production and not of well-being…
It is also flawed

Some flaws of GDP measures:

- Production and GDP vs. GNP
- How to capture changes in quality, quantity and relative prices

GDP vs. GNP
Production, Yes, …but who benefits?

- Production is still important – it is linked to employment

- GDP vs GNP:
  - privatizing resource extraction tended to generate some employment locally (hence some GDP increases) but the profits all accrue to foreign companies (seen only in GNP).
  - When you add resource depletion and environmental effects, the citizens of the county could be actually worse off…

Quality, Quantity and Relative Prices

- It is very complex to capture quality changes and price changes : e.g. electronics, computers, cars…
- Capturing change in quality as compared to quantity is a tremendous challenge:
- Example: Your mobile phone price has gone down but it can do so much more than the old phone… thus the number of units produced (where) and sold (at a particular price) is not really measuring the same thing over time.

Case Study: Mobile phones:

- **Your mobile phone price has gone down but it can do so much more than the old phone… thus the number of units produced (where) and sold (at a particular price) is not really measuring the same thing over time.**

The Modern smart mobile phone:
All of that fits in your pocket!

20 Years later and all of these things fits in you pocket.

The problem of services

- The share of services in GDP is growing in every country
- It is very difficult to measure services:
  - The problems of quality, quantity and pricing of services: medical services, ICT, educational services…
  - Public vs. private provision of services
  - Unpriced services that do not show up in accounts

How to value services?

To the extent that you have a market clearing mechanism, you can say that the amounts paid by the public for the services reflect its true worth in that society.

Measuring Services

- **The mix between public and private provision of services (education and health, housing, public sports facilities) are all valued positively by citizens.**

Measuring Services

- **Inputs vs. Outputs: Valued outputs are traditionally measured by the inputs used to produce them (e.g. number of doctors, number of hospital beds), rather than outputs (e.g. number of (successful) procedures undertaken, or number of patients treated).**

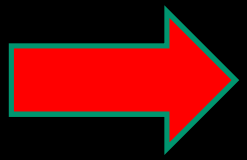
Government Services

- Growing everywhere: Government services in OECD countries have gone from about 25% of GDP to ca. 45% in the last 50 years (p.xxii)

- Generally set to approximately The wage bill of government employees
  - Absurdity of that definition
  - Consequences not followed (e.g. doubling the salaries of the civil servants)

Unremunerated labor

- Huge problem: Unremunerated labor, especially for women in the home
- Example: the Housewife vs. cook, maid, baby-sitter, housekeeper

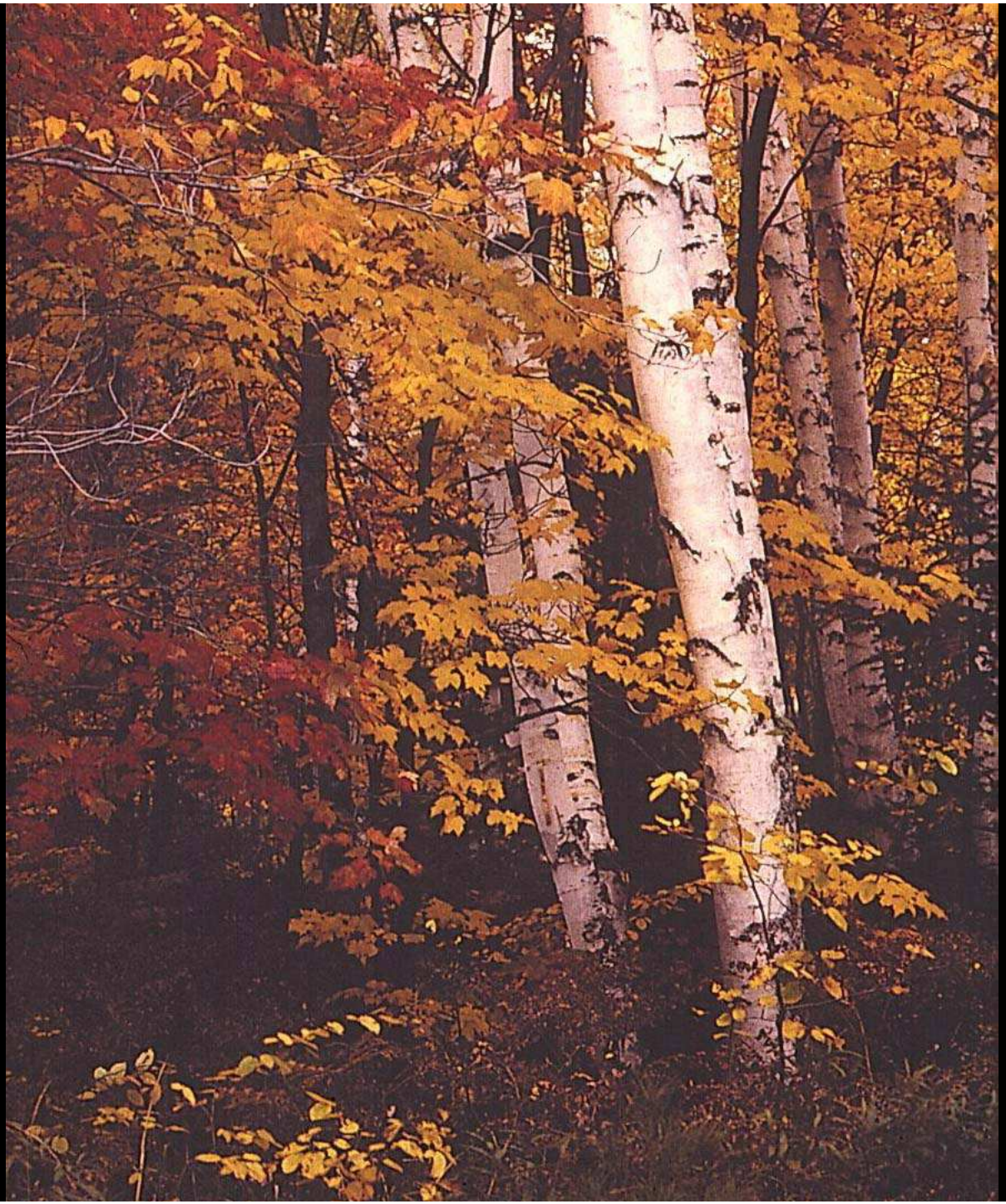


= GNP ↘

= GNP ↗

National Income Accounts

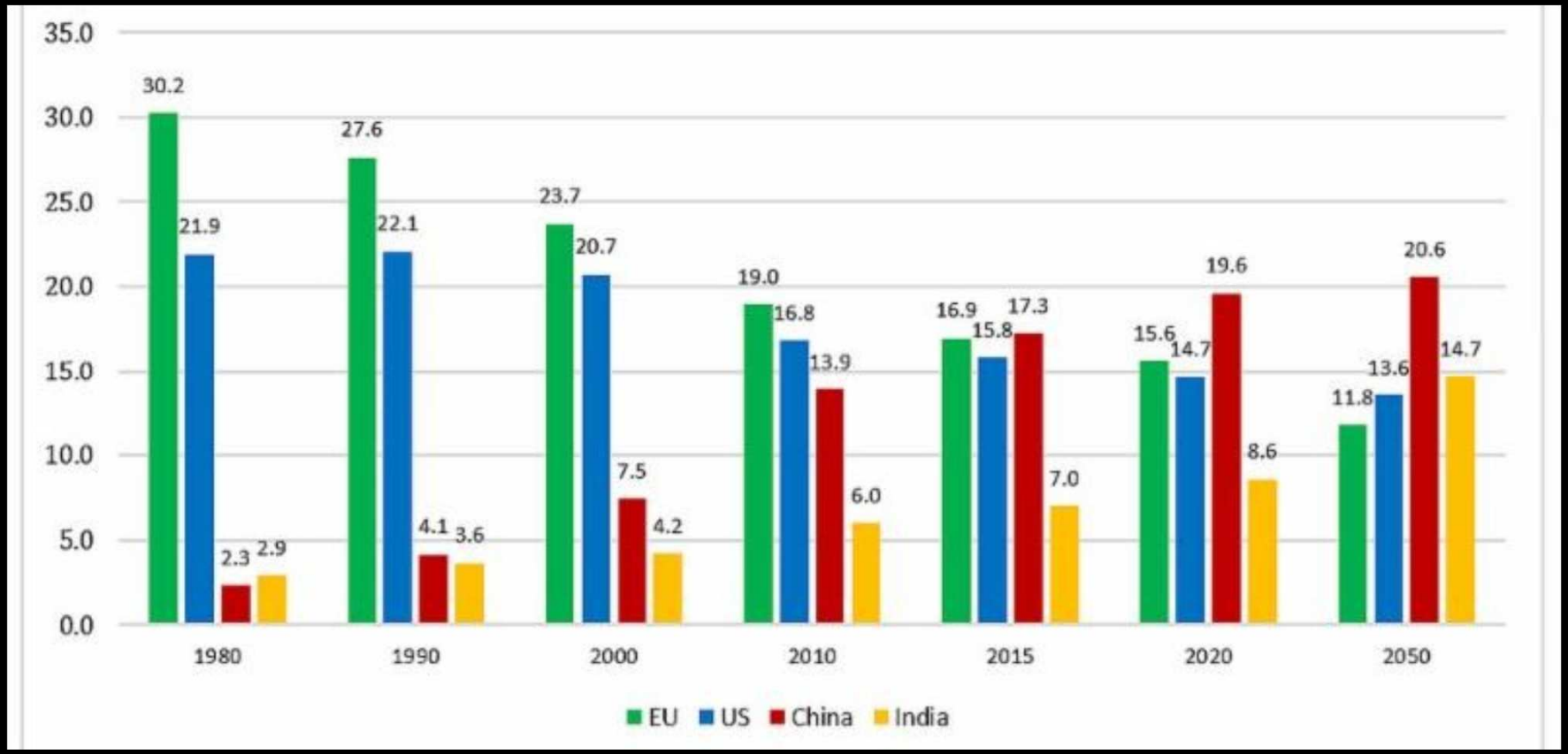
- Measure flows not stocks
- Accordingly can count a depletion of natural capital as a positive contribution
- Need to add environmental dimension
- UN agreed to add environmental accounts as satellite accounts



Country Comparisons

- GNP/GDP is used to rank size of economies and also when in per capita terms to rank by how rich the citizens are.
- But it makes a lot of difference if you use exchange rate measures or PPP.

Percentage share of global GDP (in PPP terms)

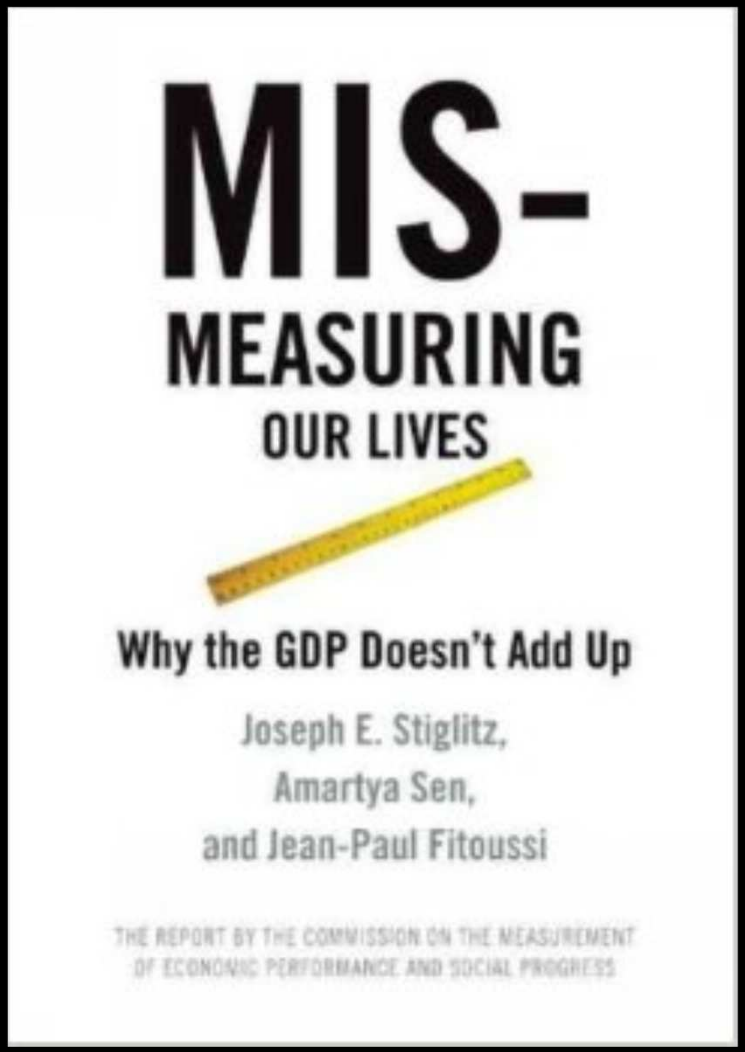


From Production to Well-being

- **More generally: It is time to shift from measuring economic production to measuring people's well-being.**

- **But because no single measure can capture well-being of people we will need a whole range of measures.**

A Truly Thoughtful Critique of GNP/GDP Measures



An Excellent Report

12 recommendations for GDP

- **Recommendation 1:** When measuring material well-being, look at income and consumption rather than production.

- **Recommendation 2:** Emphasize the household perspective

Source: Joseph E. Stiglitz, Amartya Sen and Jean-Paul Fitoussi, *Mis-Measuring Our Lives: Why GDP Doesn't Add Up* The New Press, New York, 2010, pp. xx

12 recommendations for GDP

- **Recommendation 3:** **Consider income and consumption jointly with wealth—look at the four kinds of wealth.**

Source: Joseph E. Stiglitz, Amartya Sen and Jean-Paul Fitoussi, *Mis-Measuring Our Lives: Why GDP Doesn't Add Up* The New Press, New York, 2010, pp. xx

Man-made Capital
(Produced Assets)


Natural Capital

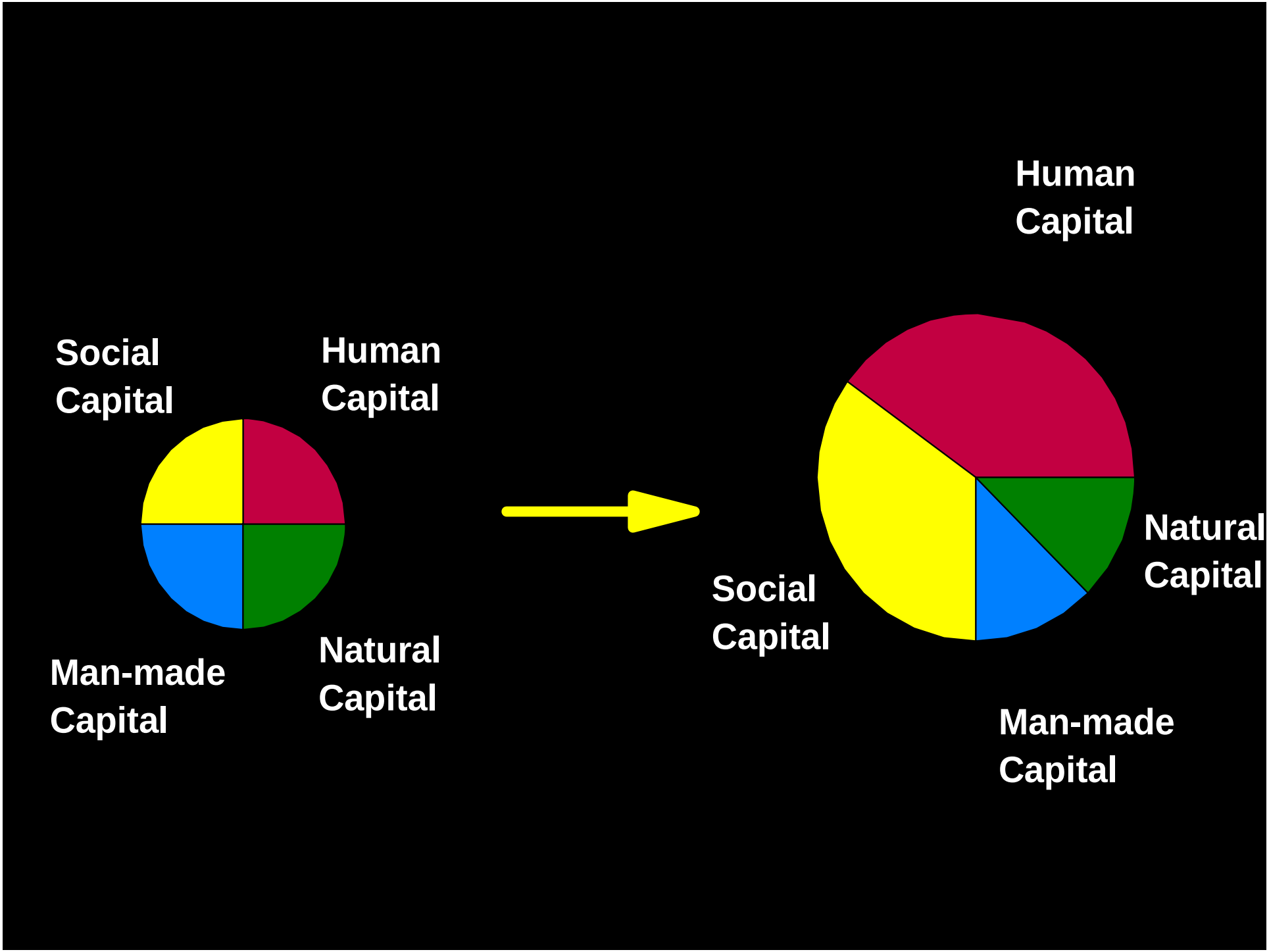

Human Capital

Social Capital

Capital per person

Comprises four kinds of capital:

- **Man-made (produced assets)**
- **Natural**
- **Human**
- **Social**


Social
Capital
Human
Capital
Man-made
Capital
Natural
Capital
Human
Capital
Social
Capital
Natural
Capital
Man-made
Capital

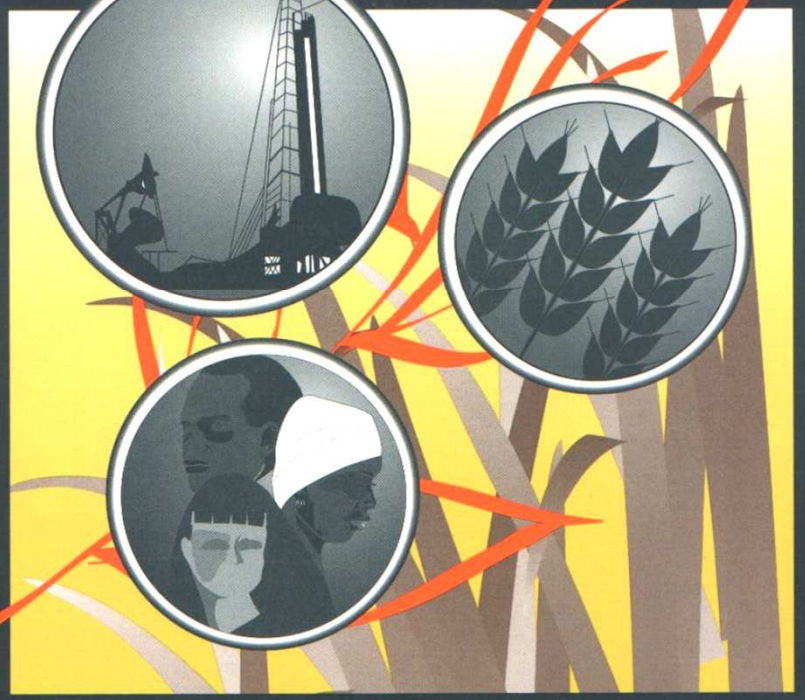
- **The four kinds of capital are partially substitutes and partially complements**

- **Therefore, mix can change over time but critical boundaries must be respected for each type of capital separately**

Sustainability and the Wealth of Nations

First Steps in an Ongoing Journey

Ismail Serageldin



ESD

Environmentally Sustainable Development Studies and Monographs Series No. 5

Wealth Accounting And Sustainability as Opportunity

12 recommendations for GDP

- **Recommendation 4:** **Give more prominence to the distribution of income, consumption and wealth**

Source: Joseph E. Stiglitz, Amartya Sen and Jean-Paul Fitoussi, *Mis-Measuring Our Lives: Why GDP Doesn't Add Up* The New Press, New York, 2010, pp. xx

12 recommendations for GDP

- **Recommendation 5:** **Broaden income measure to non-market activities (home activities, etc.) BUT ALSO Leisure: if the same amount is produced with 1500 hours rather than 2000 hours that must be an impact!**

Source: Joseph E. Stiglitz, Amartya Sen and Jean-Paul Fitoussi, *Mis-Measuring Our Lives: Why GDP Doesn't Add Up* The New Press, New York, 2010, pp. xx

To measure Well-being…

- At least eight aspects should be assessed and considered simultaneously:
  - Material living standards (income, wealth and expenditure/consumption)
  - Health
  - Education
  - Personal activities including work
  - Political voice and governance
  - Social connections and relationships
  - Environment (present and future conditions)
  - Insecurity of an economic as well as a physical nature

12 Recommendations for GDP

- **Recommendation 6: Quality of life depend on peoples' objective conditions and capabilities. Steps should be taken to improve measures of peoples' health, education, personal activities, and environmental conditions. In particular, substantial effort should be devoted to developing robust and reliable measures of social connections, political voice and insecurity that can be proven to be good predictors of life satisfaction.**

Source: Joseph E. Stiglitz, Amartya Sen and Jean-Paul Fitoussi, *Mis-Measuring Our Lives: Why GDP Doesn't Add Up* The New Press, New York, 2010, pp. xx

12 Recommendations for GDP

- **Recommendation 7: Quality of life indicators in all dimensions should also assess inequalities in a comprehensive and systematic way.**

Source: Joseph E. Stiglitz, Amartya Sen and Jean-Paul Fitoussi, *Mis-Measuring Our Lives: Why GDP Doesn't Add Up* The New Press, New York, 2010, pp. xx

12 Recommendations for GDP

- **Recommendation 8:** **surveys should be designed to assess the links between various quality of life domains for each person, and this information should be used when designing policies in various fields.**

Source: Joseph E. Stiglitz, Amartya Sen and Jean-Paul Fitoussi, *Mis-Measuring Our Lives: Why GDP Doesn't Add Up* The New Press, New York, 2010, pp. xx

12 Recommendations for GDP

- **Recommendation 9:** **Statistical offices should provide the information needed to aggregate across quality of life dimensions and thereby allowing the construction of various indices.**

Source: Joseph E. Stiglitz, Amartya Sen and Jean-Paul Fitoussi, *Mis-Measuring Our Lives: Why GDP Doesn't Add Up* The New Press, New York, 2010, pp. xx

12 Recommendations for GDP

- **Recommendation 10:** **Measures of both objective and subjective well-being provide information for people's quality of life. Statistical offices should incorporate questions to capture people's life evaluations, hedonic experiences and priorities in their own surveys.**

Source: Joseph E. Stiglitz, Amartya Sen and Jean-Paul Fitoussi, *Mis-Measuring Our Lives: Why GDP Doesn't Add Up* The New Press, New York, 2010, pp. xx

12 Recommendations for GDP

- **Recommendation 11:** **Sustainability assessment requires a well-identified dashboard of indicators… using stocks… hence going back to sustainability as opportunity and the four kinds of capital (see Serageldin).**

Source: Joseph E. Stiglitz, Amartya Sen and Jean-Paul Fitoussi, *Mis-Measuring Our Lives: Why GDP Doesn't Add Up* The New Press, New York, 2010, pp. xx

12 Recommendations for GDP

- **Recommendation 12:** **environmental indicators of sustainability should be based on physical indicators and also include tipping point analysis (proximity to dangerous levels of environmental damage, climate change, collapse of fish stocks, etc.).**

Source: Joseph E. Stiglitz, Amartya Sen and Jean-Paul Fitoussi, *Mis-Measuring Our Lives: Why GDP Doesn't Add Up* The New Press, New York, 2010, pp. xx

Measuring Poverty

Measuring Poverty

- Headcount Index
- Depth of Poverty (Poverty Gap)
- Foster-Greer-Thorbecke Index (P_α)

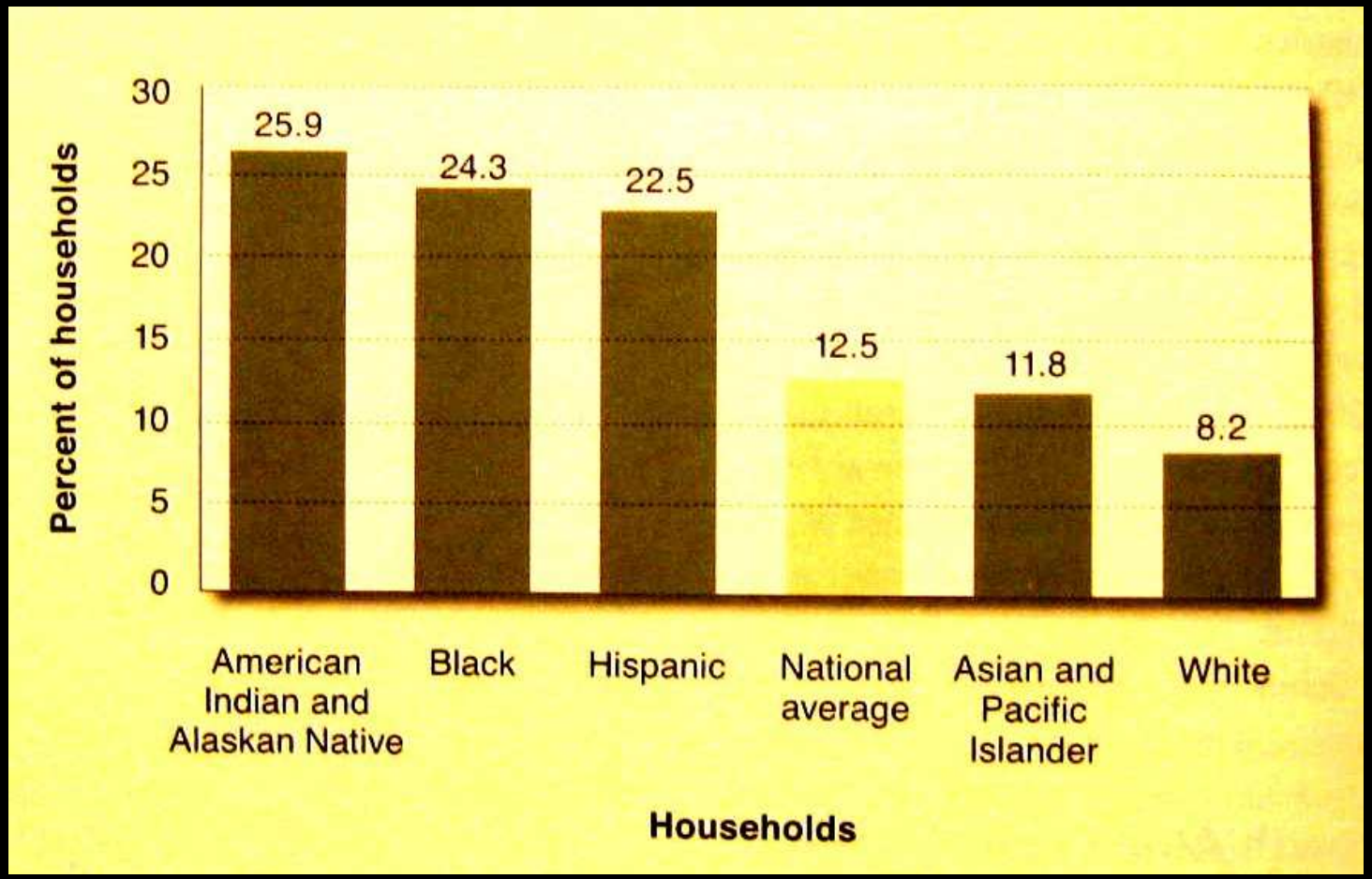
Headcount Index

The Headcount Index is the proportion of people below the poverty line:

$$H = \frac{q}{n}$$

where n is total population and q is population whose Y < z and z is the Y at Poverty Line

USA % Households below the poverty line (2003)



Depth of Poverty (Poverty Gap)

Measures how far the average poor person is below the poverty line and multiplies that by the headcount Index

$$P_G = \frac{1}{n}\sum_{i=1}^{q}\left[\frac{z - y_i}{z}\right]$$

$\therefore \quad P_G = I.H$ where $I = \frac{Z - y^i}{Z}$

and I = mean depth of poverty as a proportion of the poverty Line

P_G= Cost of eliminating poverty by Y- transfer to the poor.

A	α	alpha	N	ν	nu
B	β	beta	Ξ	ξ	xi
Γ	γ	gamma	O	o	omicron
Δ	δ	delta	Π	π	pi
E	ϵ	epsilon	P	ρ	rho
Z	ζ	zeta	Σ	σ	sigma
H	η	eta	T	τ	tau
Θ	θ	theta	Υ	υ	upsilon
I	ι	iota	Φ	ϕ	phi
K	κ	kappa	X	χ	chi
Λ	λ	lambda	Ψ	ψ	psi
M	μ	mu	Ω	ω	omega

Greek Alphabet

P_α : The FGT Poverty Index

$$P_\alpha = \frac{1}{n}\sum_{i=1}^{q}\left[\frac{z - y_i}{z}\right]^\alpha$$

P_α : The FGT Poverty Index

$$P_\alpha = \frac{1}{n} \sum_{i=1}^{q} \left[\frac{z - y_i}{z} \right]^\alpha$$

P_α

- If $\alpha = 0 \quad \therefore P_0 =$ Headcount Index
- If $\alpha = 1 \quad \therefore P_1 =$ Poverty Gap Measure
- If $\alpha = 2 \quad \therefore P_2 =$ Mean of squared proportionate poverty gaps

A better statement about P_α

- If $\alpha = 0 \quad \therefore P_0 =$ **Amount** of poverty
- If $\alpha = 1 \quad \therefore P_1 =$ **Depth** of Poverty
- If $\alpha = 2 \quad \therefore P_2 =$ **Severity** of Poverty (usually associated with hunger)

Understanding P_α

$$\therefore P_\alpha = \frac{1}{n}\sum_{i=1}^{q}\left[\frac{z - y_i}{z}\right]^\alpha$$

P_α is the weighted mean over the poor population

The measure = $\left(1 - {y_i}/{z}\right)^\alpha$ for poor $(y_i < z)$

= 0 for non $-$ poor $(y_i > z)$

P_2 -- what does it measure?

If α = 2, then P_2= mean of squared proportionate poverty gaps … AND:

$$P_2 = \underbrace{\frac{PG^2}{H}}_{\text{Contribution of pov.gap to } P_2} + \underbrace{\frac{(H-PG)^2}{H}.CV_P^2}_{\text{Contribution of inequality amongst the poor to } P_2}$$

The FGT P_α Indexes

- Provide much richer measurement

The FGT P_α Indexes

- Provide much richer measurement
- But P_2 is difficult to interpret for decision makers

The FGT P_α Indexes

- Provide much richer measurement
- But P_2 is difficult to interpret for decision makers
- **Decomposable**

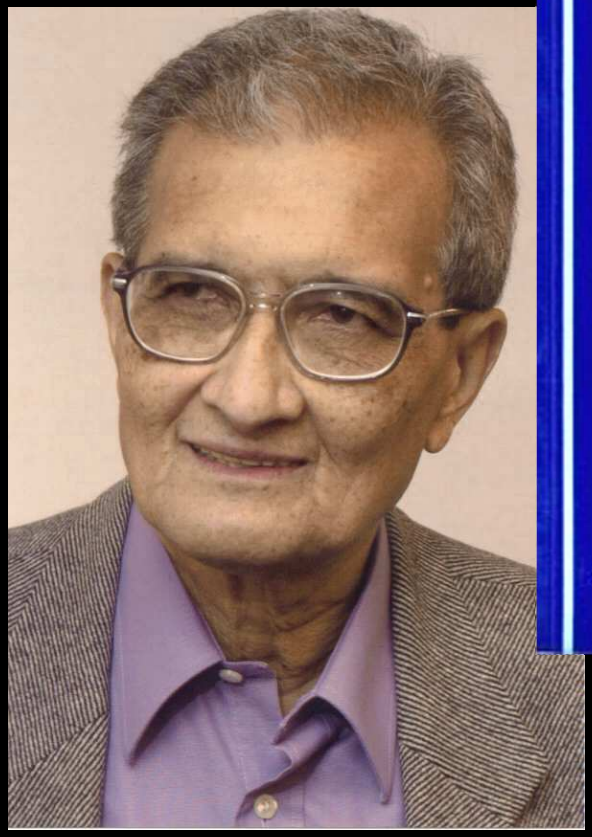
The FGT P_α Indexes

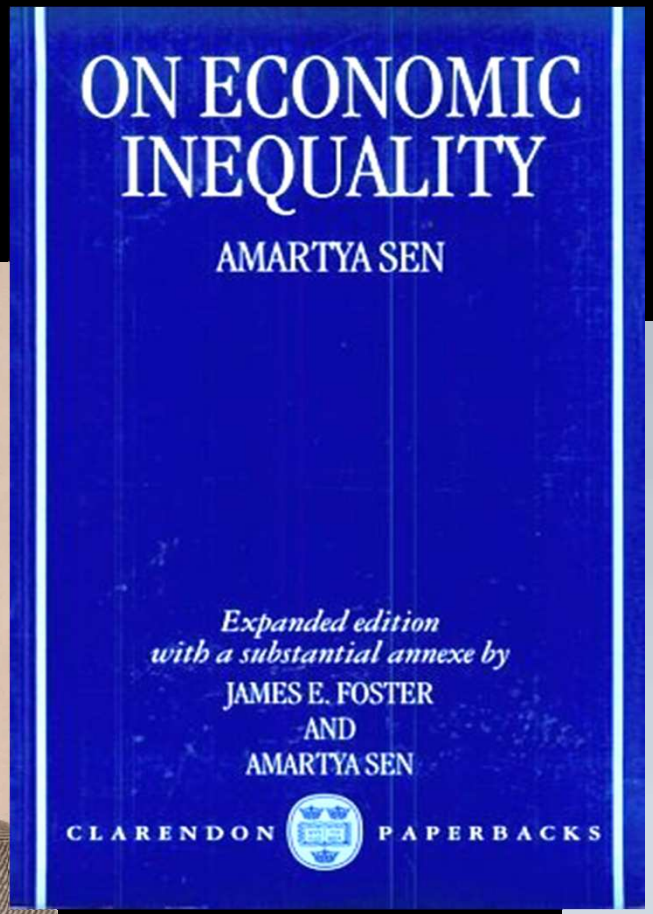
- Provide much richer measurement
- But P_2 is difficult to interpret for decision makers
- **Decomposable – an attractive feature, BUT…**

The underlying mathematical structure for the index allows partitioning the set (decomposability) by any dimension, no matter how absurd

This means that ultimately it is an individual measure and voids the relational or social context aspect of poverty.

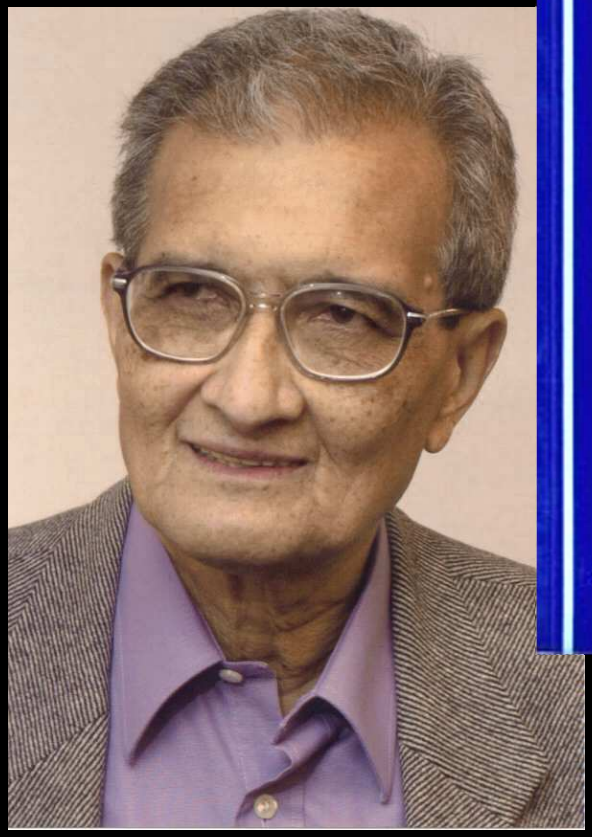
Ideally, the use of FGT indexes should be supplemented by inequality indicators.

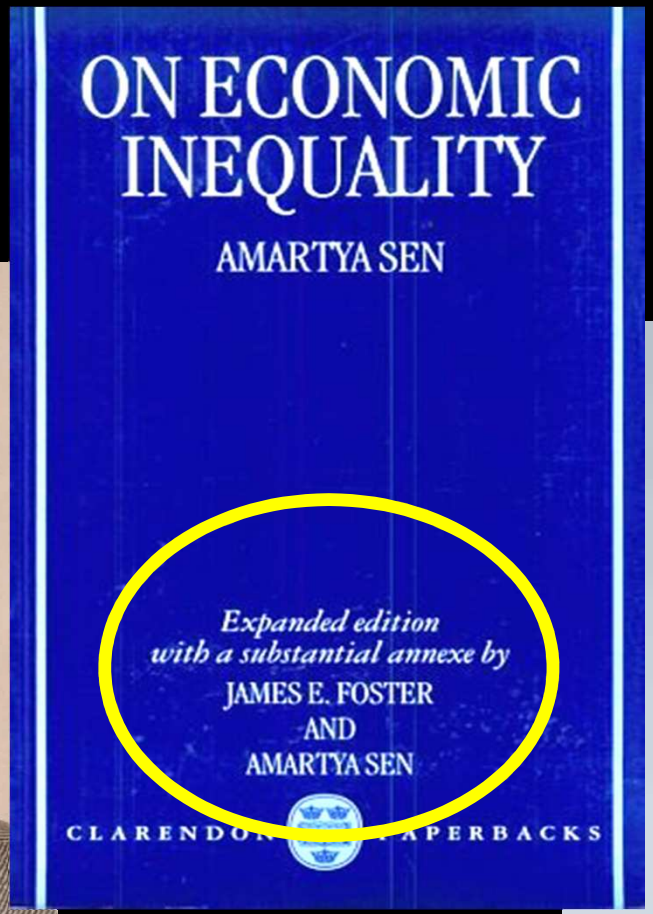




Amartya Sen

James Foster



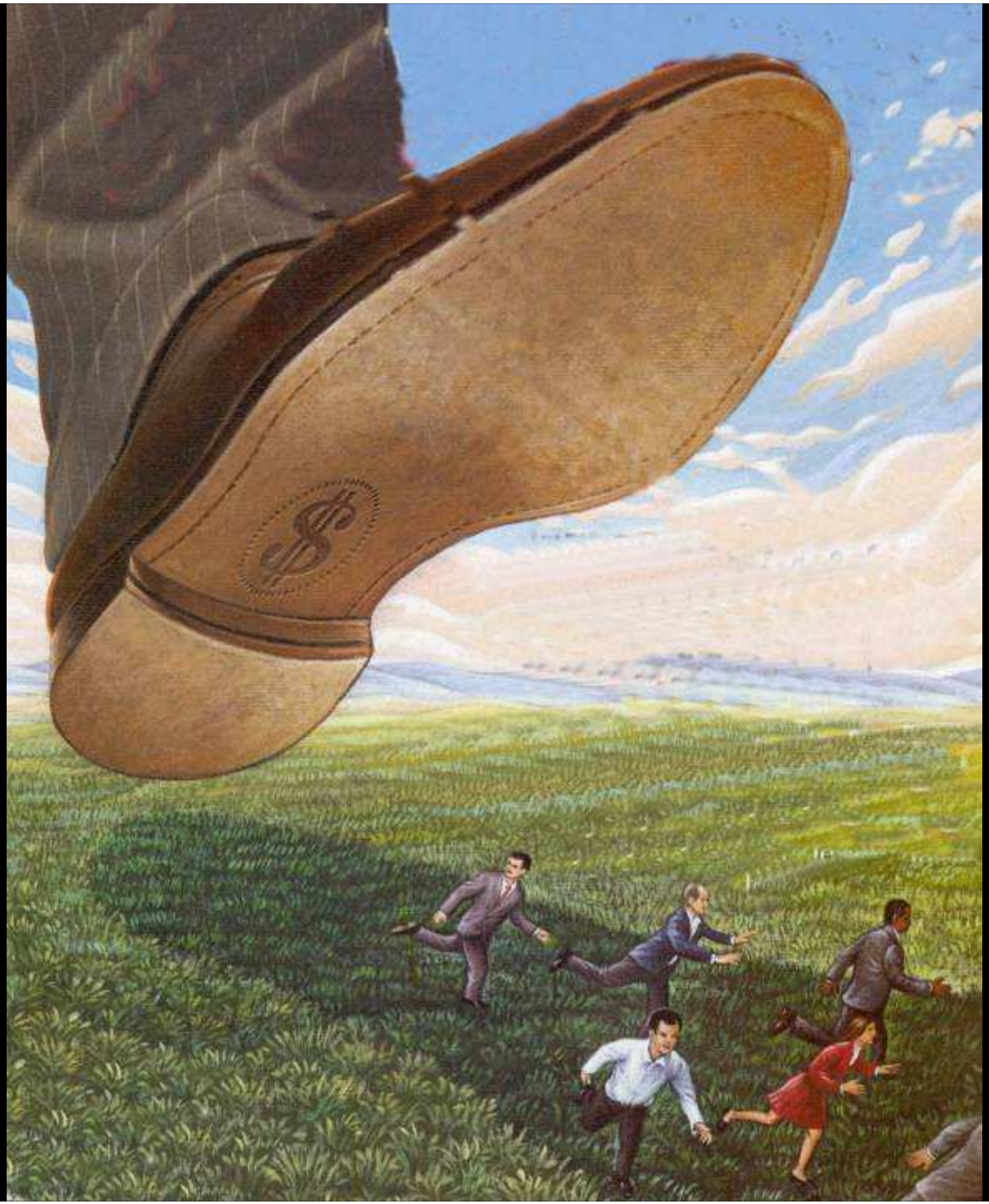


Amartya Sen

James Foster

Links To Inequality

Measuring Inequality

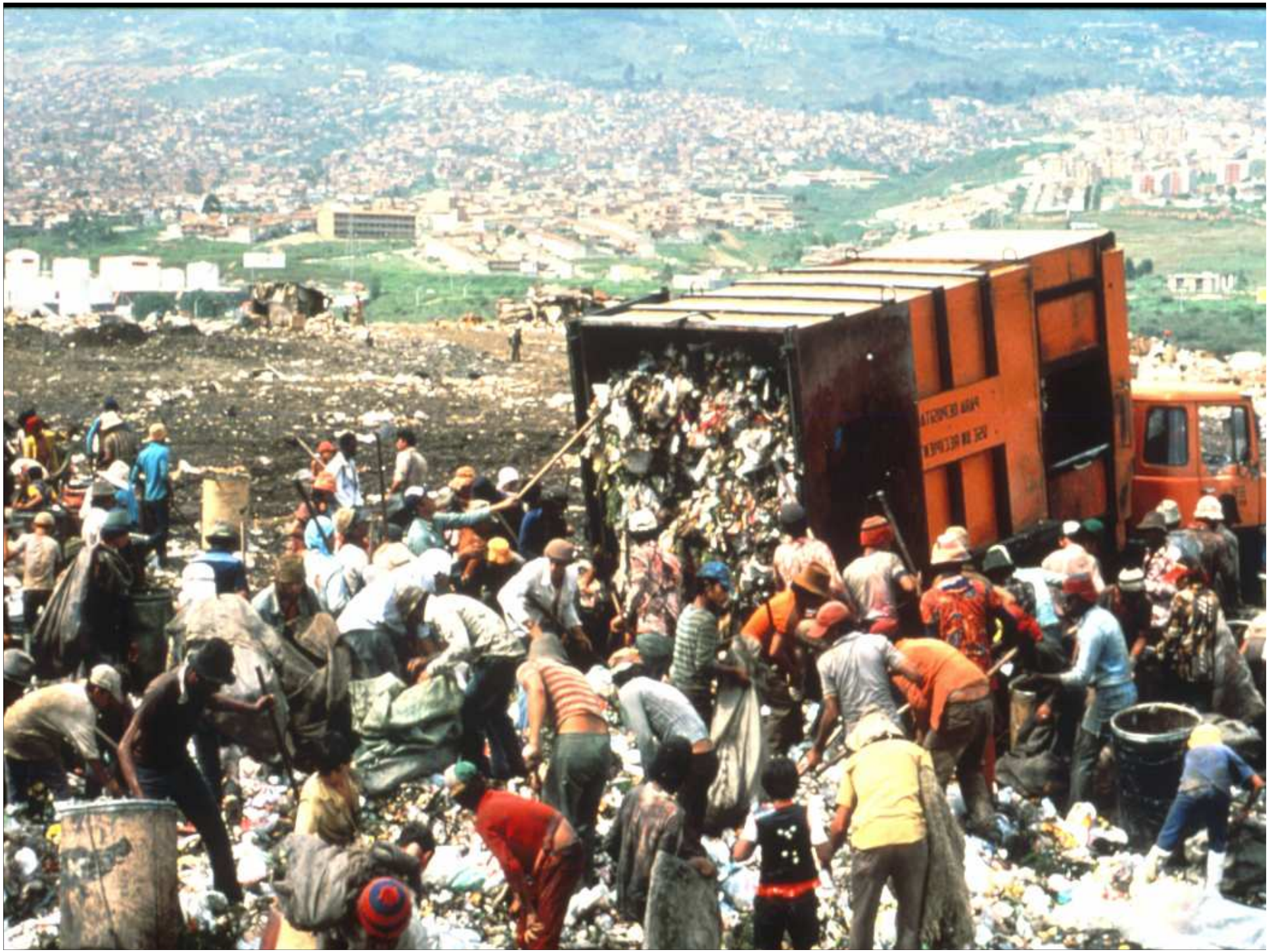


An Enormous Gap Exists Between the Rich and the Poor…


TIME


†
I AM
HOMELESS
PLEASE HELP


5 24 16 33 1 20 14 31 9
11 30 8 23
0



Rich and Poor in Sao Paulo

source: http://mindblog.dericbownds.net/2007/10/rich-and-poor.html

The Most Widely Accepted and Used Measures of Inequality

- The Gini Coefficient and the Lorenz Curve
- Closely interrelated
- Powerfully descriptive

Origins

- The Gini coefficient was developed by the Italian Statistician Corrado Gini (Gini, 1912) as a summary measure of income inequality in society.
- It is usually associated with the plot of wealth concentration introduced a few years earlier by Max Lorenz (Lorenz, 1905).
- Since these measures were introduced, they have been applied to topics other than income and wealth, but mostly within Economics (Cowell, 1995, 2000; Jenkins, 1991; Sen, 1973).

Max Otto Lorenz (1876 – 1959)

- He developed the Lorenz curve in 1905 to describe income inequalities.

- He published this paper when he was a doctoral student at the University of Wisconsin–Madison.

Corrado Gini
(1884-1965)

- Corrado Gini was an Italian statistician, demographer and sociologist

- He developed the Gini coefficient, a measure of the income inequality in a society in 1912.

Gini Coefficient

- Inequality on the Gini scale is measured between 0, where everybody is equal, and 1, where all the country's income is earned by a single person.
- It allows comparing inequality between countries or within the same country over time.

See inter alia, Sen, A. On Economic Inequality. Oxford, England: Clarendon Press, 1973. Or http://mathworld.wolfram.com/GiniCoefficient.html (Accessed 24 01 2018)

Amartya Sen: A Pairwise Comparison

- When G is based on the Lorenz curve of income distribution, it can be interpreted as the expected income gap between two individuals randomly selected from the population (Sen, 1973).

The classical definition of G appears in the notation of the theory of relative mean difference:

$$G = \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} \left| x_i - x_j \right|}{2n^2 \bar{x}}$$

Where:

x is an observed value

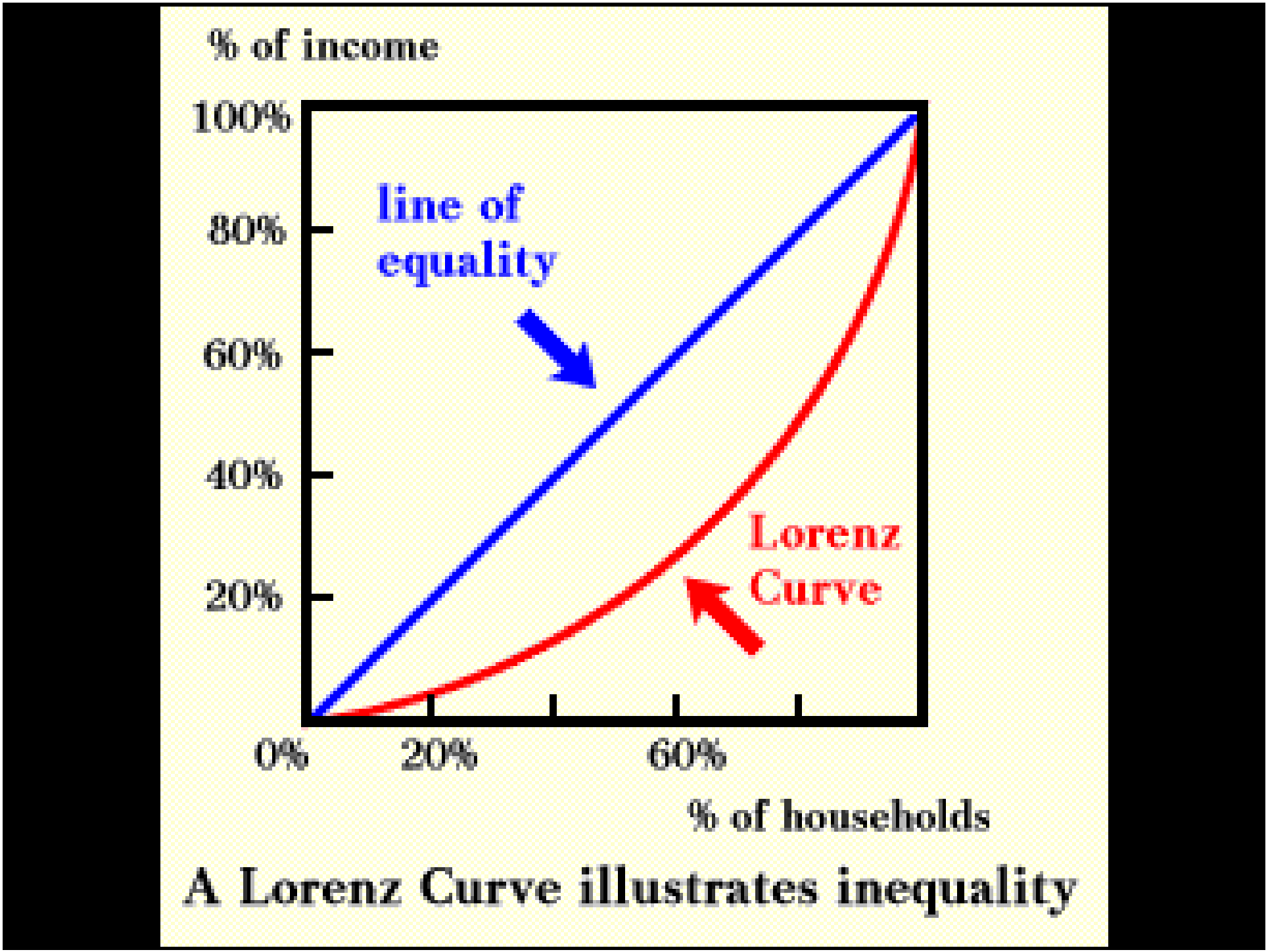
n is the number of values observed

$\bar{x}$ is the mean value

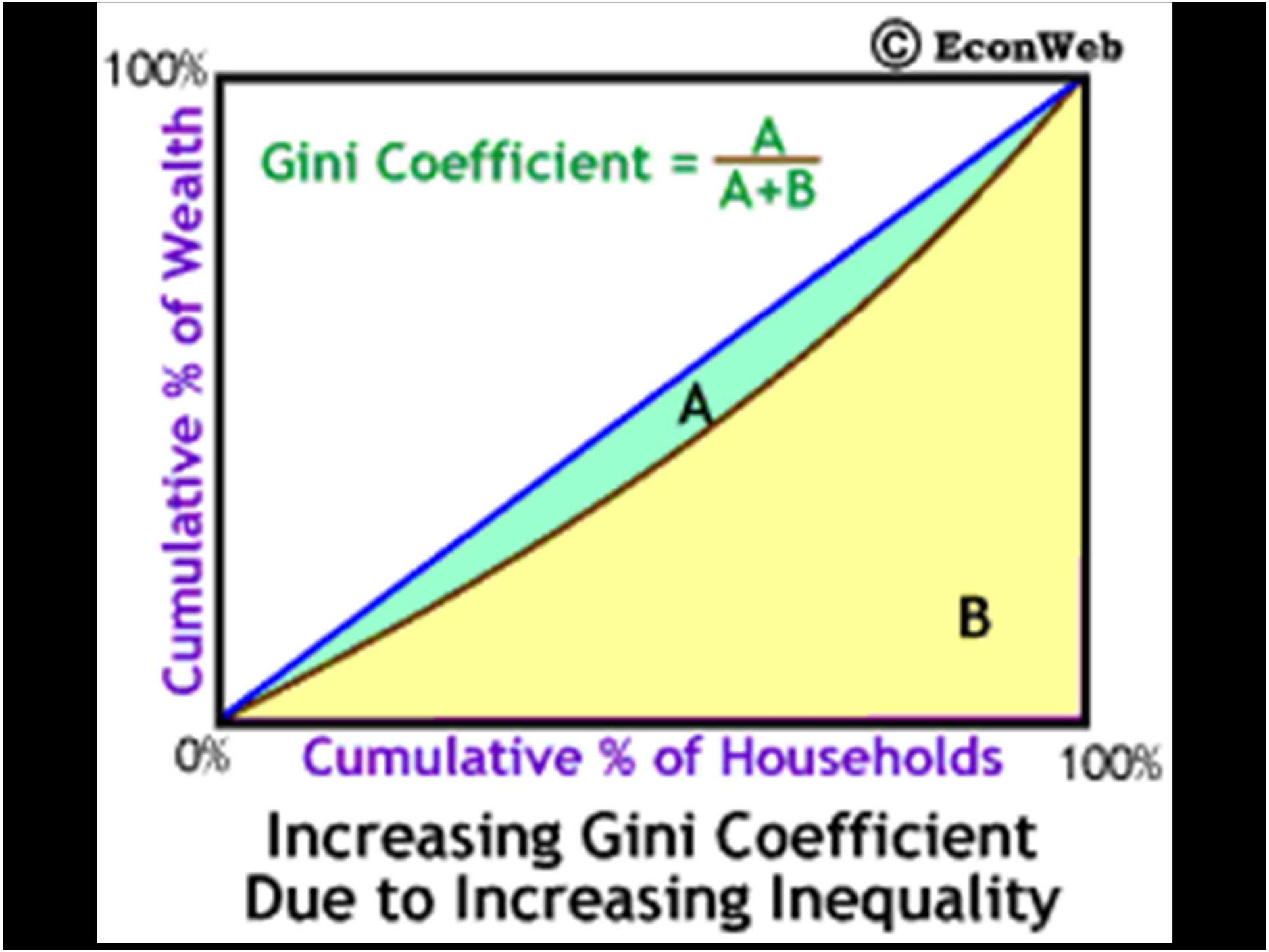
Measuring Inequality: The Lorenz Curve and the Gini Coefficient

- The Gini coefficient (or Gini ratio) is a summary statistic of the Lorenz curve and a measure of inequality in a population.
- The Gini coefficient is most easily calculated from unordered size data as the "relative mean difference," i.e., the mean of the difference between every possible pair of individuals, divided by the mean size …

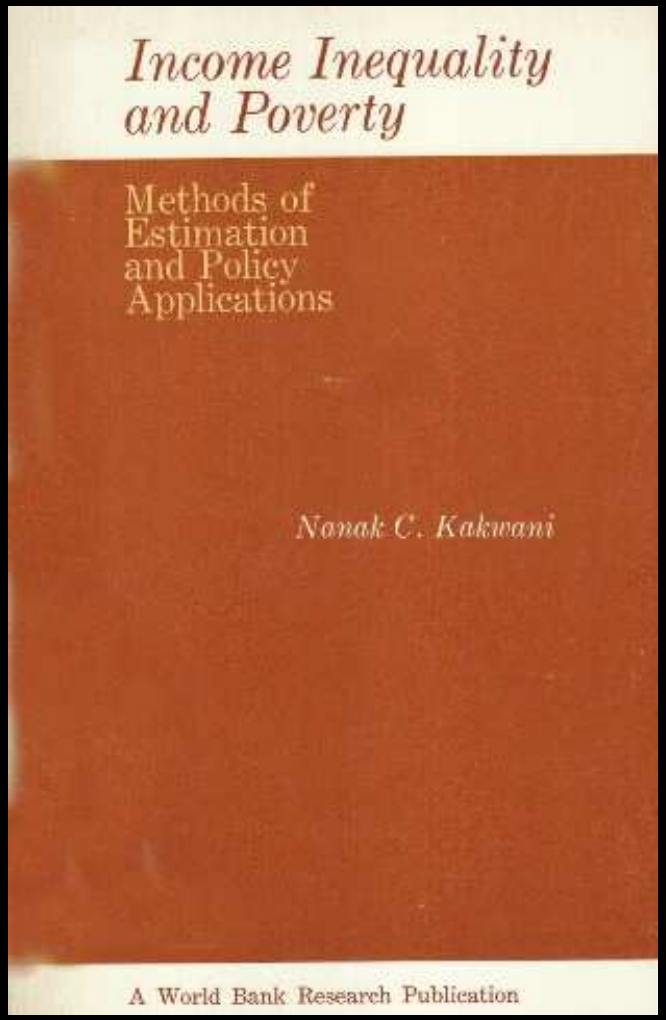

Index: 0 = perfect equality; 1 = perfect inequality
GINI COEFFICIENT FOR FAMILIES
0.46
0.44
0.42
0.40
0.38
0.36
0.34
1960
1965
1970
1975
1980
1985
1990
1995
2000



A Lorenz Curve illustrates inequality


© EconWeb
100%
Cumulative % of Wealth
Gini Coefficient = A / A+B
A
B
0%
Cumulative % of Households
100%
Increasing Gini Coefficient
Due to Increasing Inequality

Additional Measures



Kakwani (as well as others) have also sugested other ways of measuring, including comparing the length of the Lorenz Curve to the diagonal

Thoughtful use of multiple indicators for analysis will lead to richer and more nuanced policy and program design

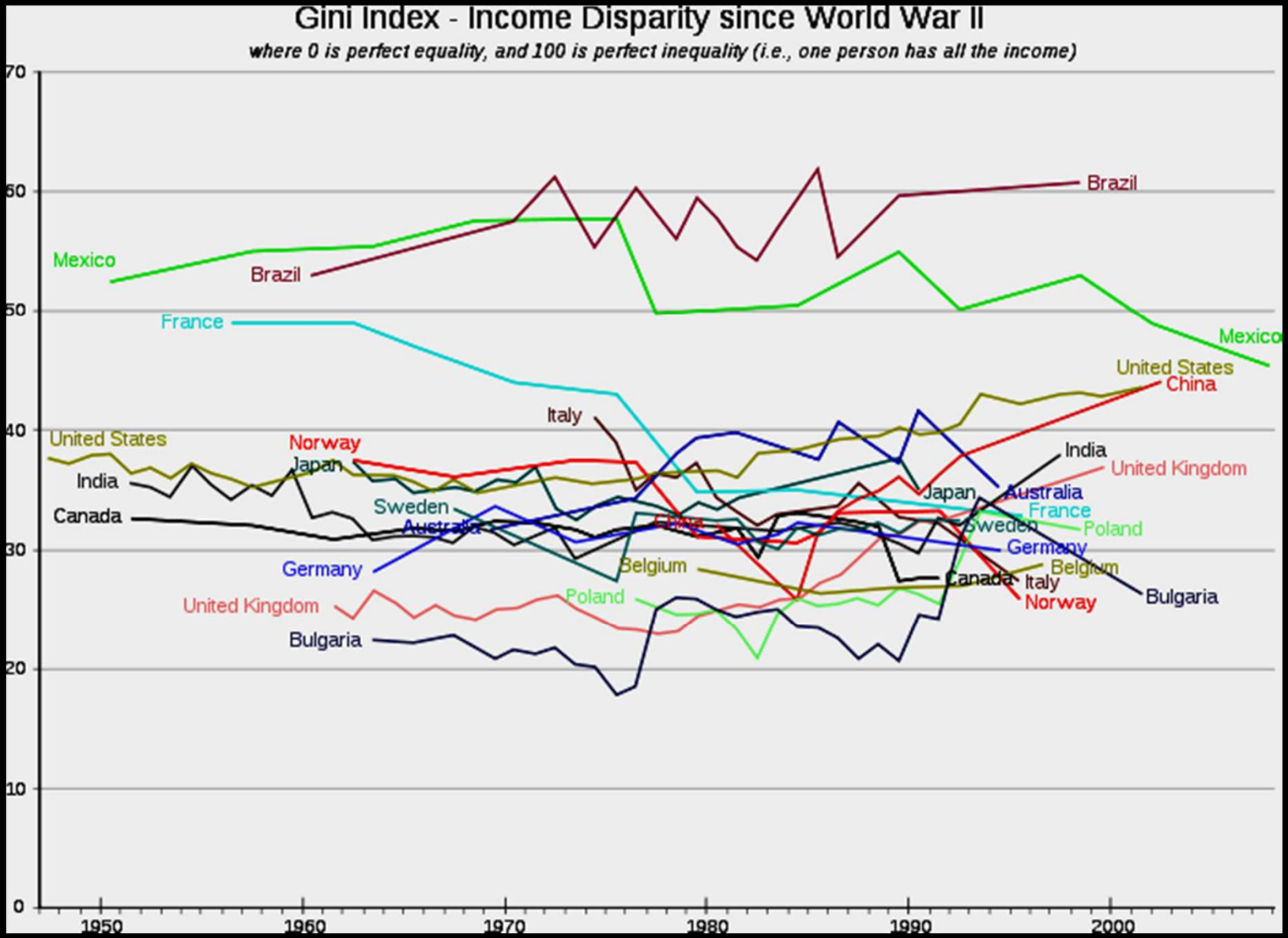
But index numbers allow shorthand indications of status, trends and inter-country comparisons
(all with due caution!)

List extracted from the Gini Index for a selected group of nations

Japan	24.9	United Kingdom	36.0
Sweden	25.0	Iran	43.0
Germany	28.3	United States	46.6
France	32.7	Argentina	52.2
Pakistan	33.0	Mexico	54.6
Canada	33.1	South Africa	57.8
Switzerland	33.1	Namibia	70.7

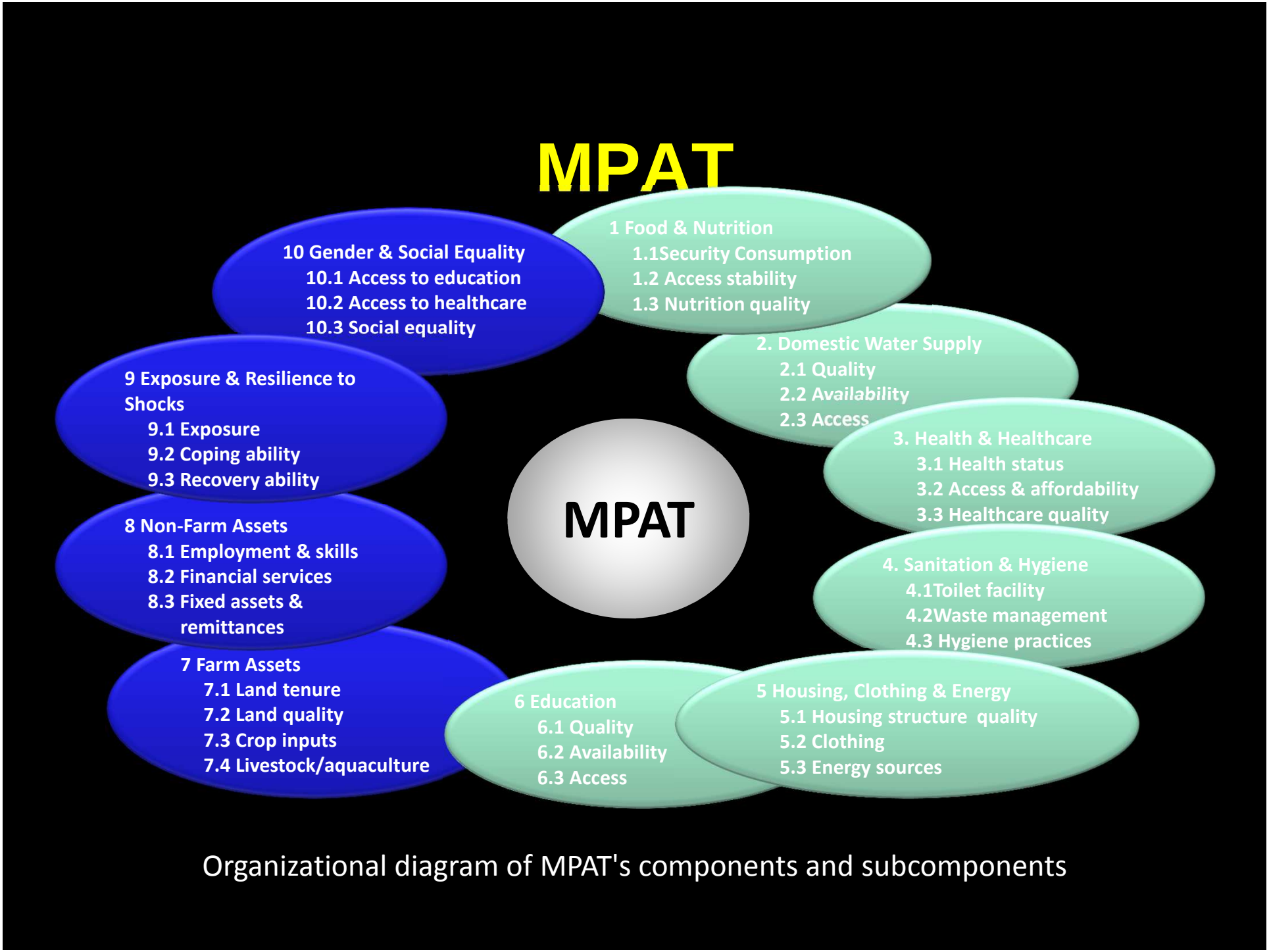

Index: 0 = perfect equality; 1 = perfect inequality
GINI COEFFICIENT FOR FAMILIES
0.46
0.44
0.42
0.40
0.38
0.36
0.34
Trend
1960
1965
1970
1975
1980
1985
1990
1995
2000

Gini Index – Income Disparity Since WWII



Thoughtful use of multiple indicators for analysis will lead to richer and more nuanced policy and program design

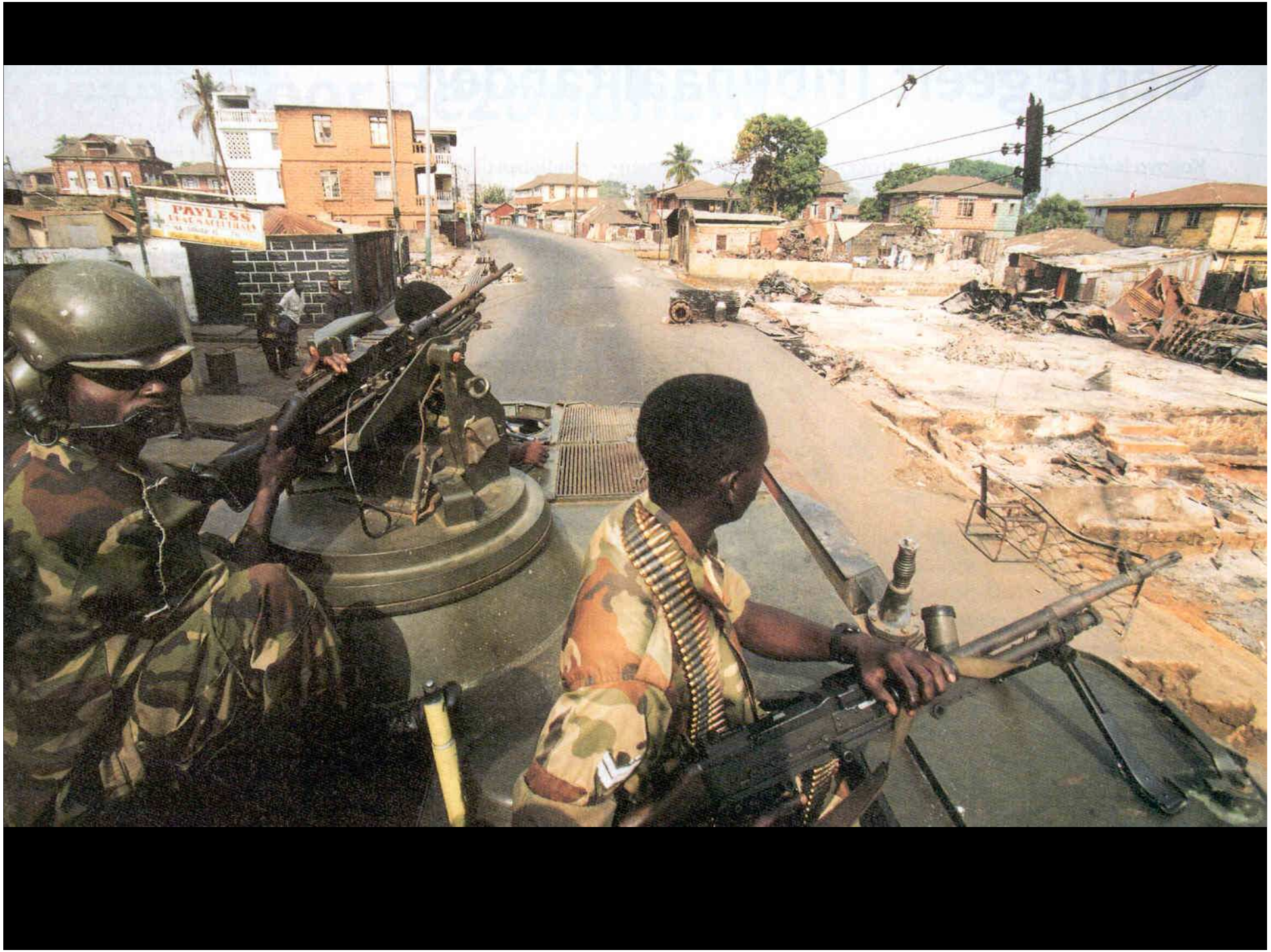
MPAT:
Multidimensional Poverty Assessment Tool
(Developed by IFAD)

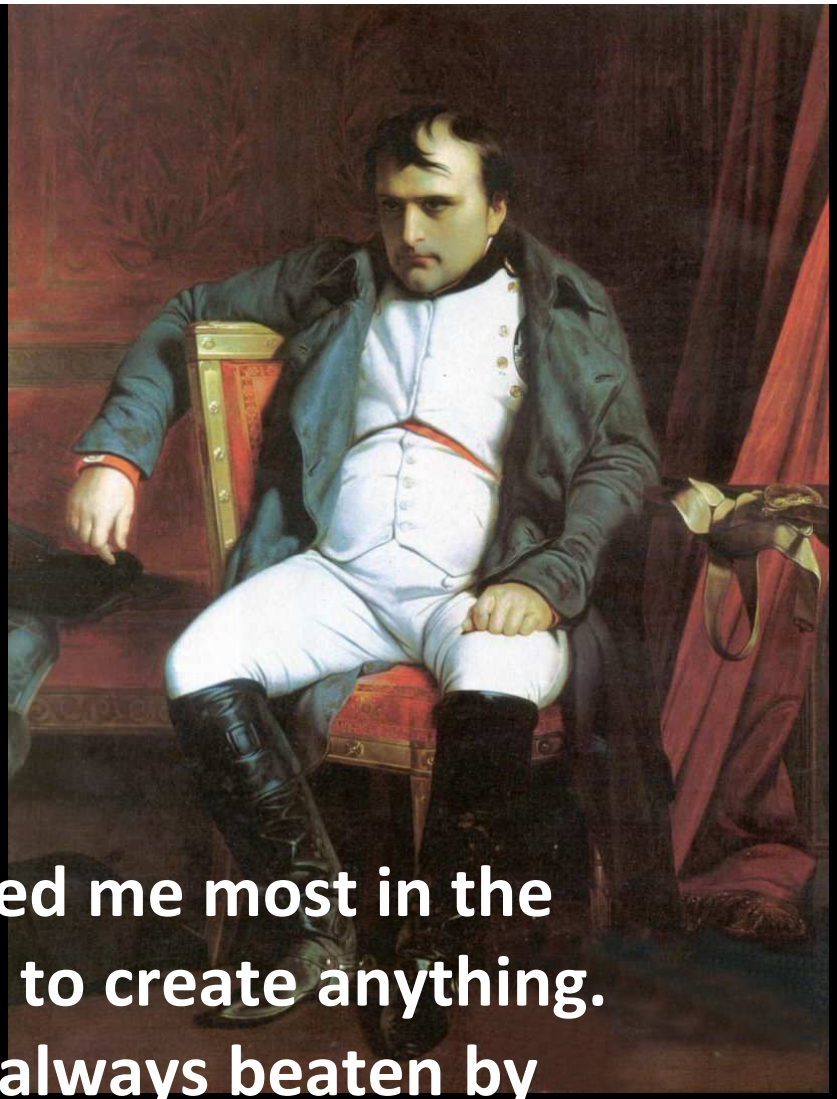


Organizational diagram of MPAT's components and subcomponents

Item # 9:
On Resilience to Shocks

Extreme Poverty and Resilience


PAYLESS

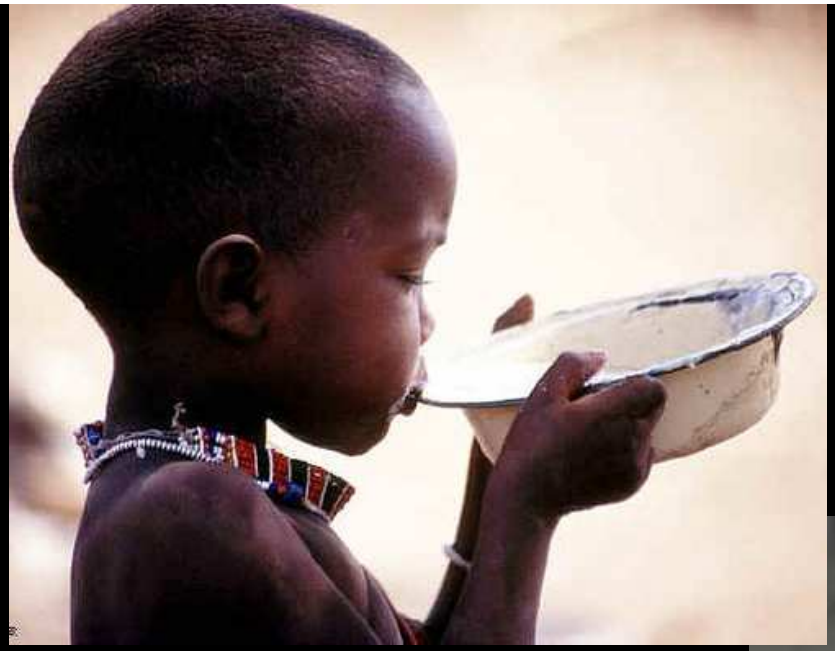


- **Do you know what astonished me most in the world? The inability of force to create anything. In the long run the sword is always beaten by the spirit.**
  - **Napoleon Bonaparte (1769-1821)**

Books Or Bombs?

The vulture and the little girl

- Also known as "Struggling Girl", attempting to reach a UN feeding center in Ayod, South Sudan in March 1993.
- This photograph by Kevin Carter first appeared in The NYT on 26 March 1993.
- Won Pulitzer Prize in 1994.
- Kevin Carter Committed suicide four months later in 1994… he was 33.



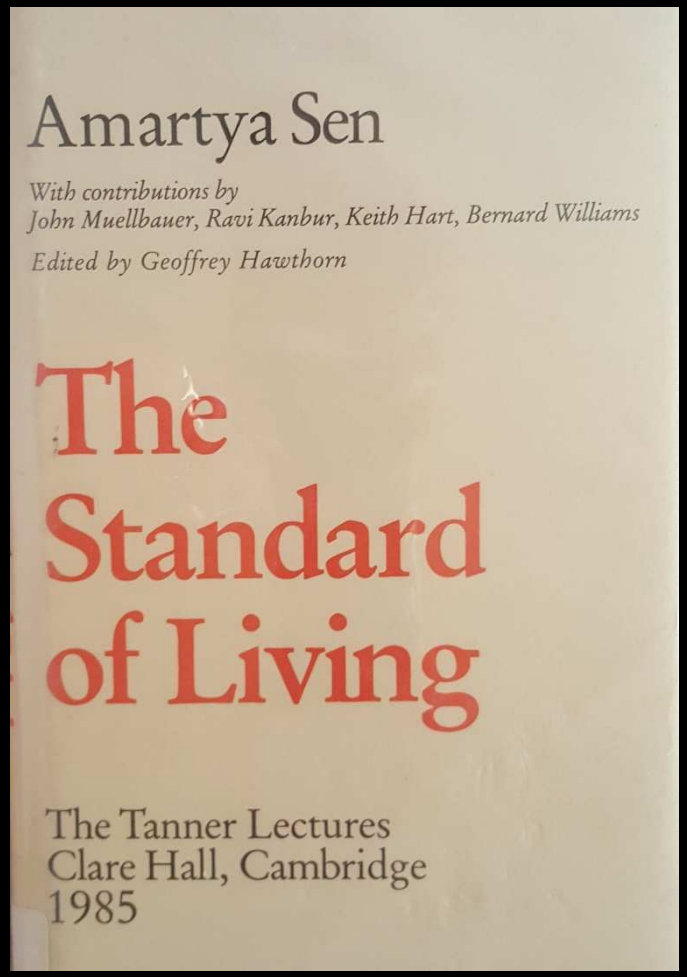
Kevin Carter
(1960 - 1994)

We need a lot more work on the development of proper techniques to understand and promote resilience in communities at risk of shocks…

Some Further Conceptual Issues

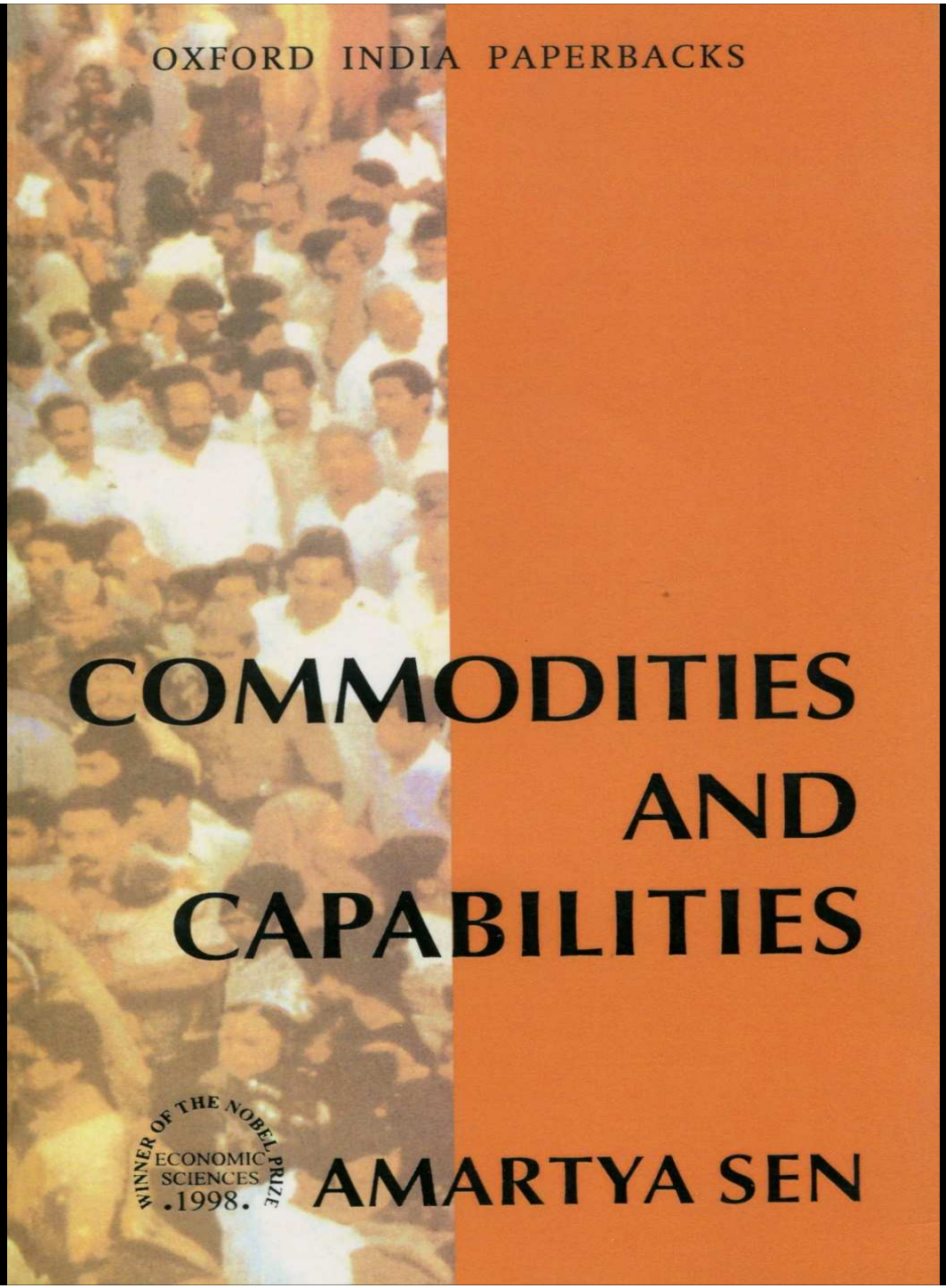
Poverty as captured by the limited capabilities of the poor

Already in 1985 …



- **The 1985 Tanner lectures in Cambridge.**
- **Contributions by John Muellbauer, Ravi Kanbur, Keith Hart, & Bernard Williams**
- **Edited by Geoffrey Hawthorn.**


Diagram by John Muellbauer


OXFORD INDIA PAPERBACKS
COMMODITIES AND CAPABILITIES
WINNER OF THE NOBEL PRIZE
ECONOMIC SCIENCES
1998
AMARTYA SEN

Some of Sen's Well-known Views Found in the Monograph

- Raises many issues on the conceptual foundations of welfare economics, diverging from the usual concepts of income or wealth (opulence) or utility.
- Looks at the challenges of thinking about what a person can do, or can be, given their capabilities and the obstacles that society puts on their ability to benefit fully of the supposed equality of opportunity in the system.

This also addresses issues of sex-bias and ethnic or religious discrimination, etc.

But lets get back to the more simple measures we have been discussing

From Data To Models

We like single numbers because:

- Shorthand way of giving an indication of the size of a problem
- Easier to compare changes over time
- Easier to compare with other cases (countries, locations, groups) where they have the same number

But what we gain in convenience we lose in diagnostic power, accuracy of interpretation, and as guidance for effective policies.

SO:

Always Try to Understand:

- How the index number is constructed
- What the number does NOT capture
- How relevant to the question at hand is that index
- How relevant are the things that it misses

Then, and only then,

- You can use such numbers with care

On
Generating the Data
We Need

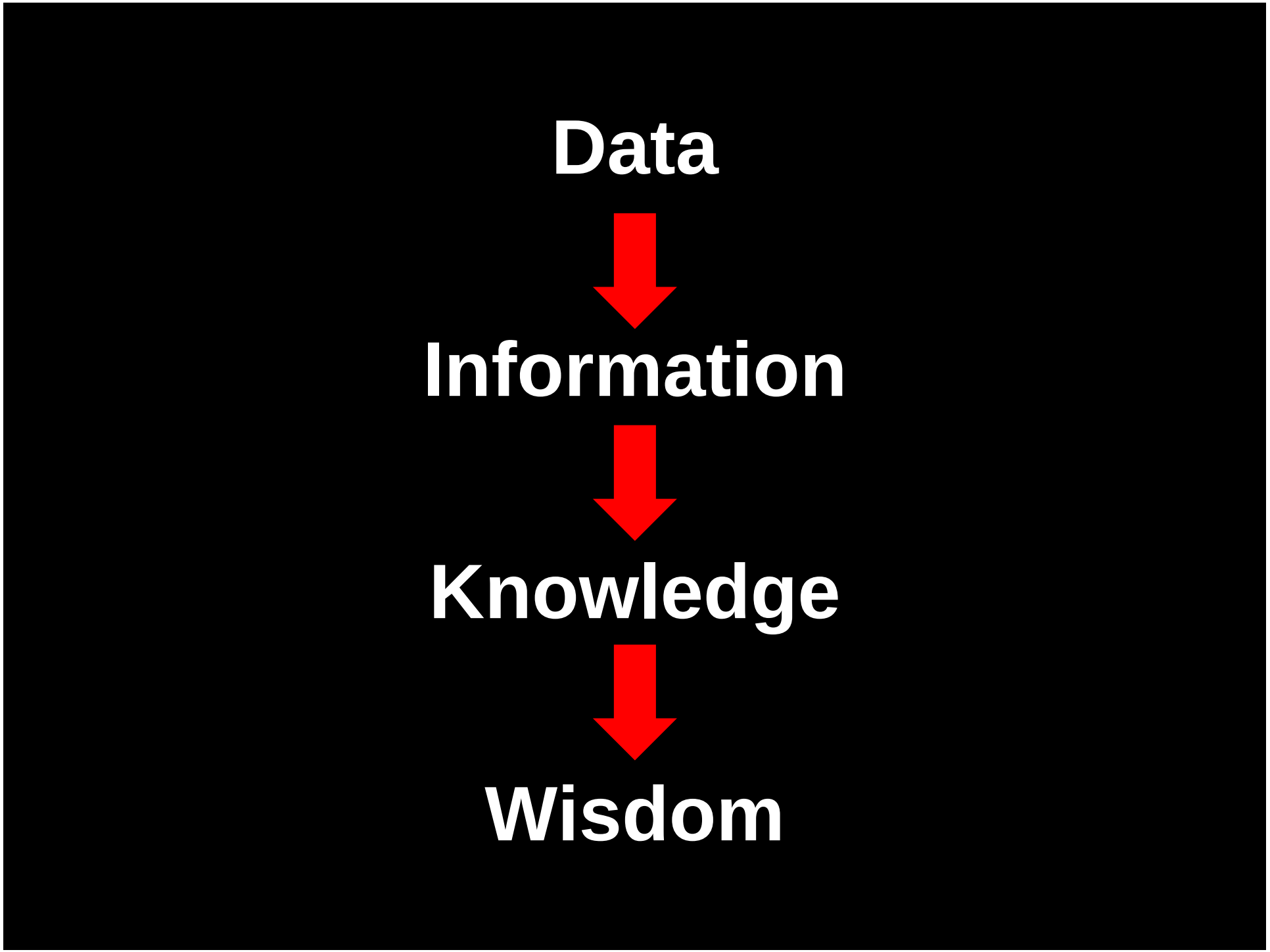
Conventional
Sources of Data and Insights

- Census;
- Sample Surveys; and
- Research design (experimental and quasi-experimental designs for research)

New
Sources of Data and Insights

- Social media;
- Earth Observation (EO); and
- AI, Machine learning, targeted discovery, etc.
- And much more…

And usually, when we have the data we build mathematical models to help us diagnose, test scenarios and guide policies


Data
Information
Knowledge
Wisdom

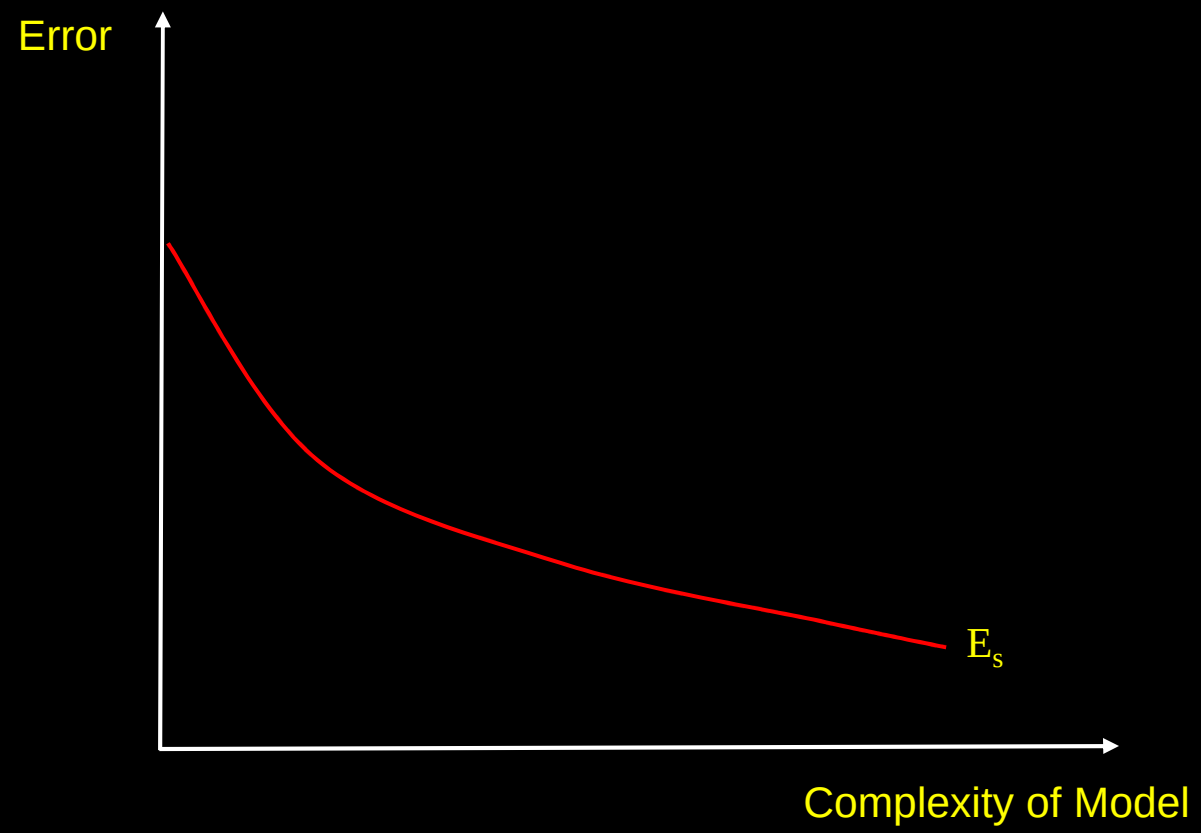
Building Models: Design And Errors

Building Mathematical Models

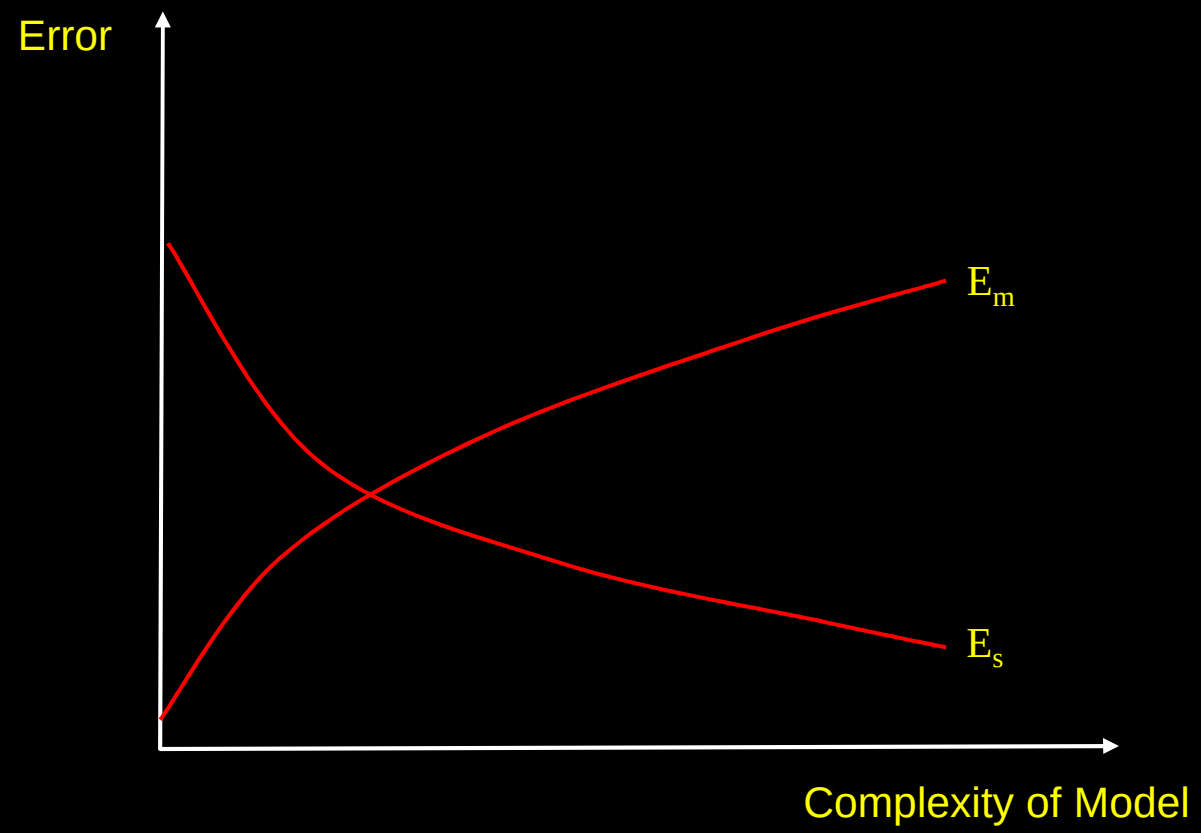
Purely descriptive models can help in basic understanding of possible causalities and identification of likely intervention points

Predictive models are needed for policy analysis. This requires quantification.

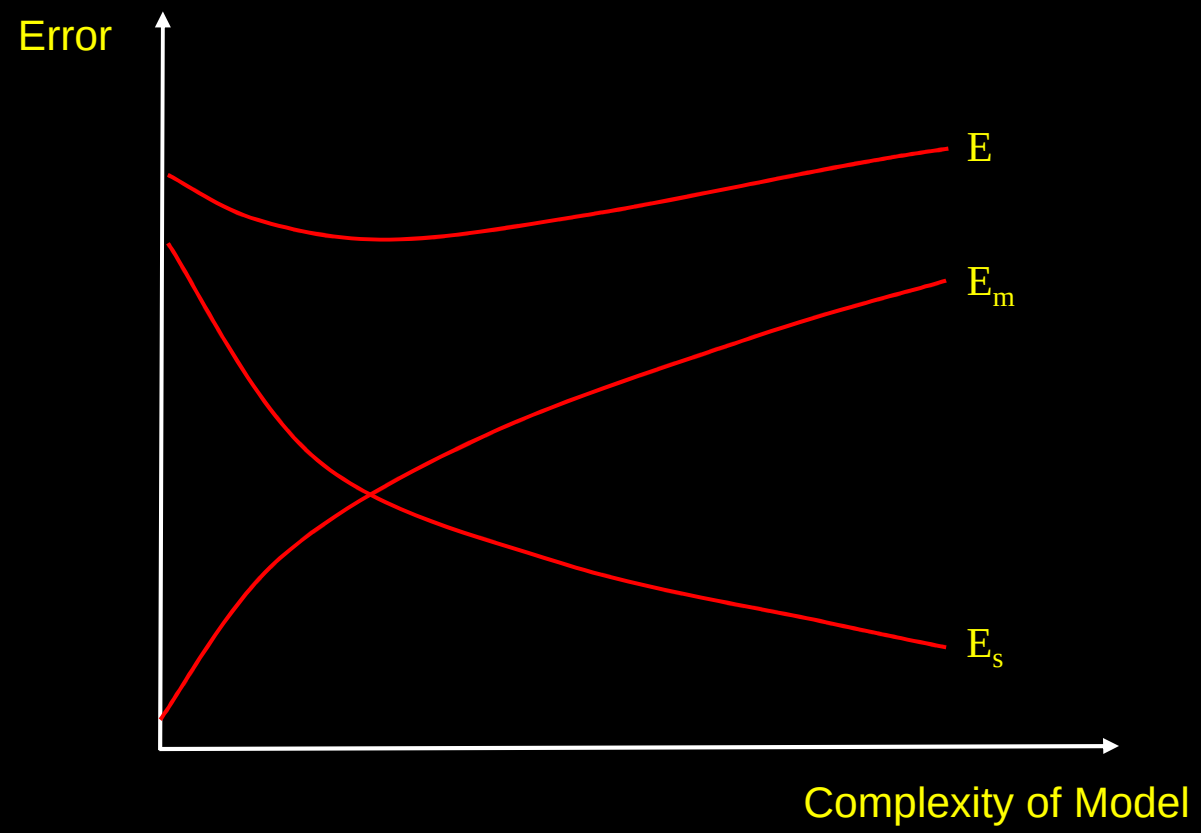
Errors in Models



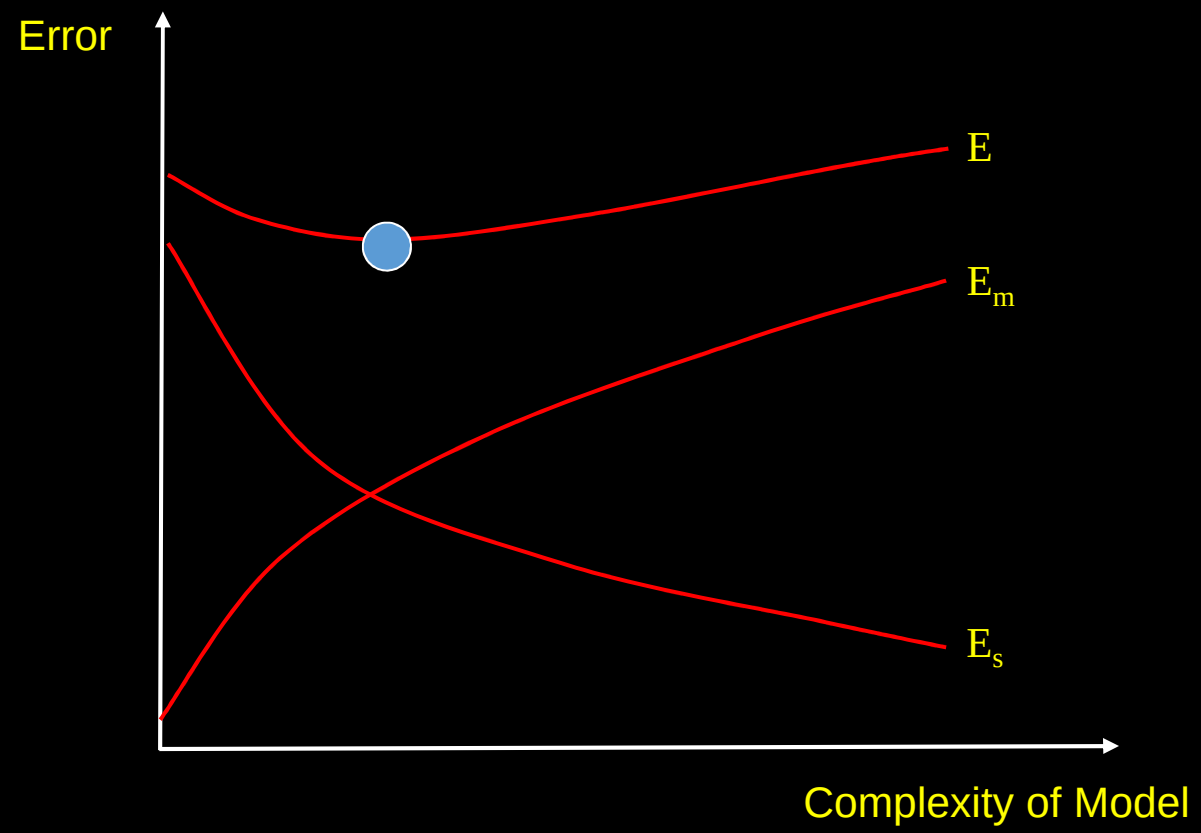
Errors in Models



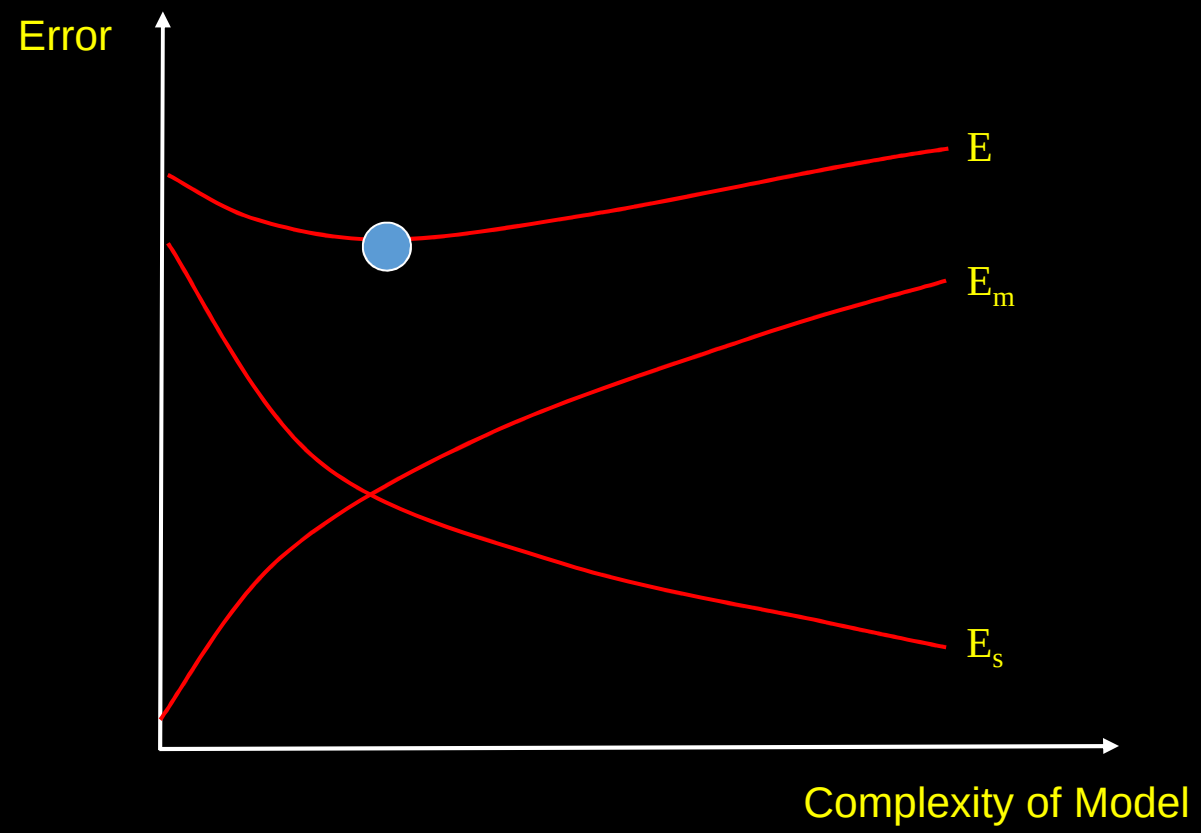
Errors in Models



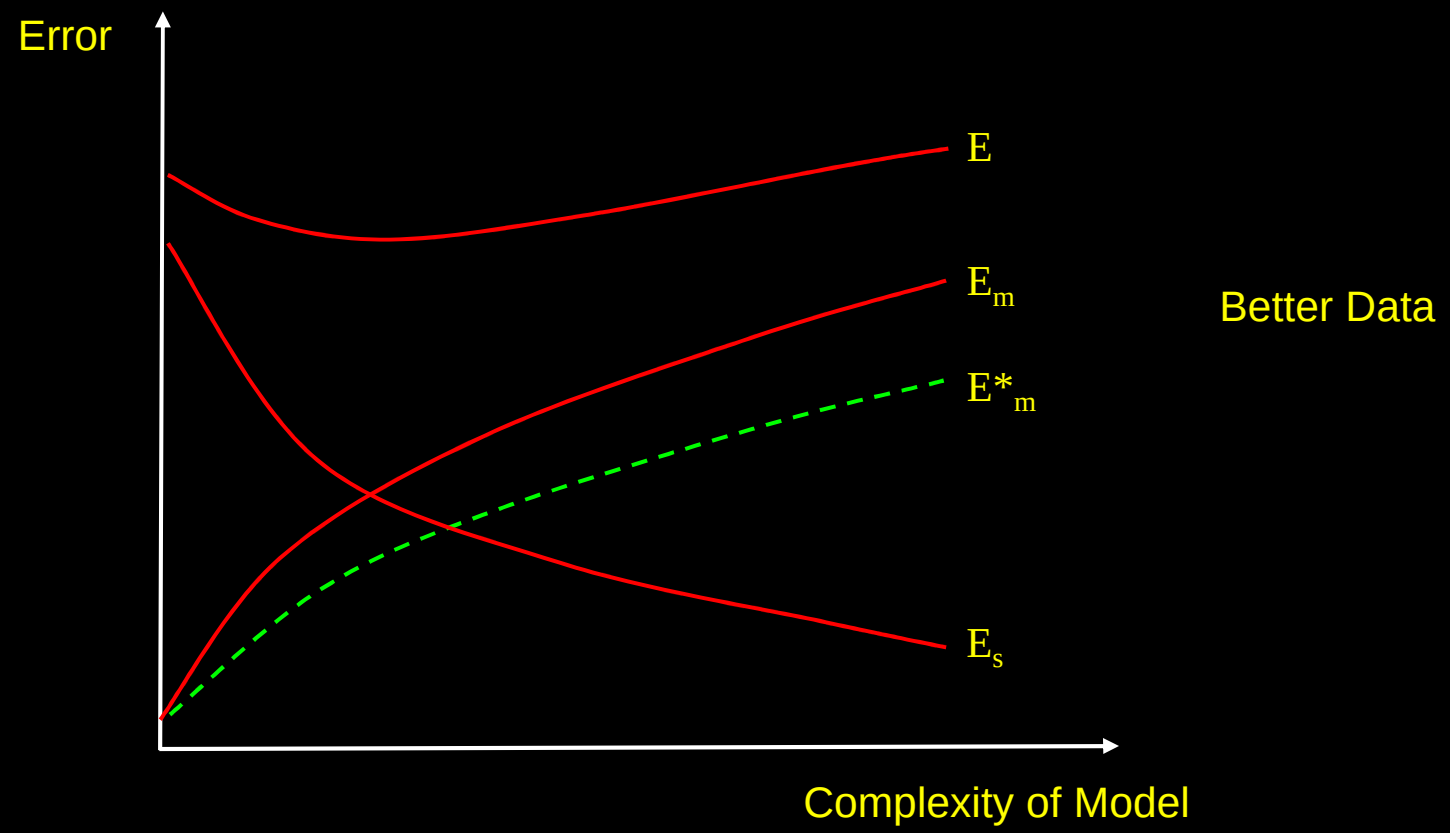
Errors in Models



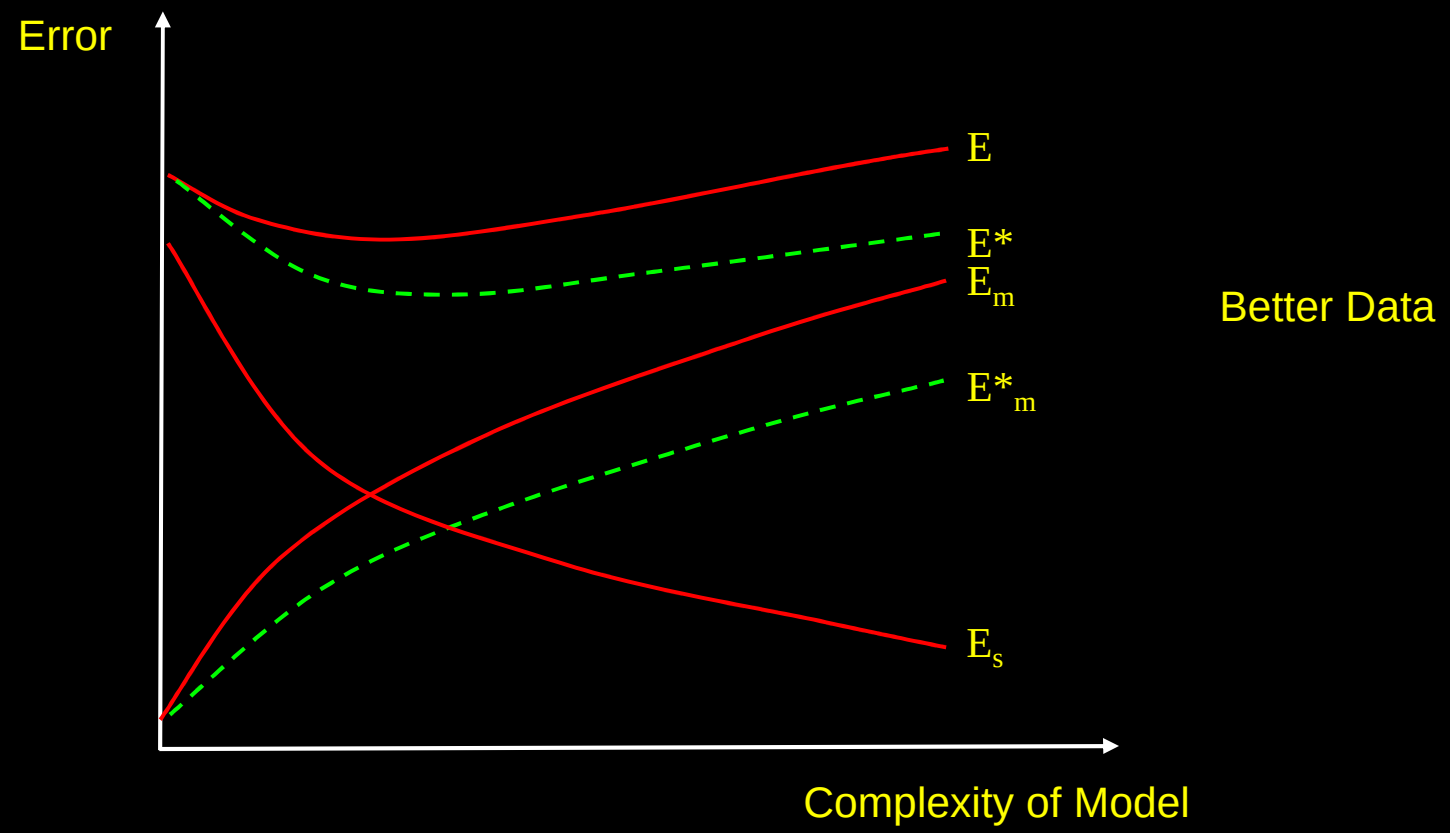
Errors in Models



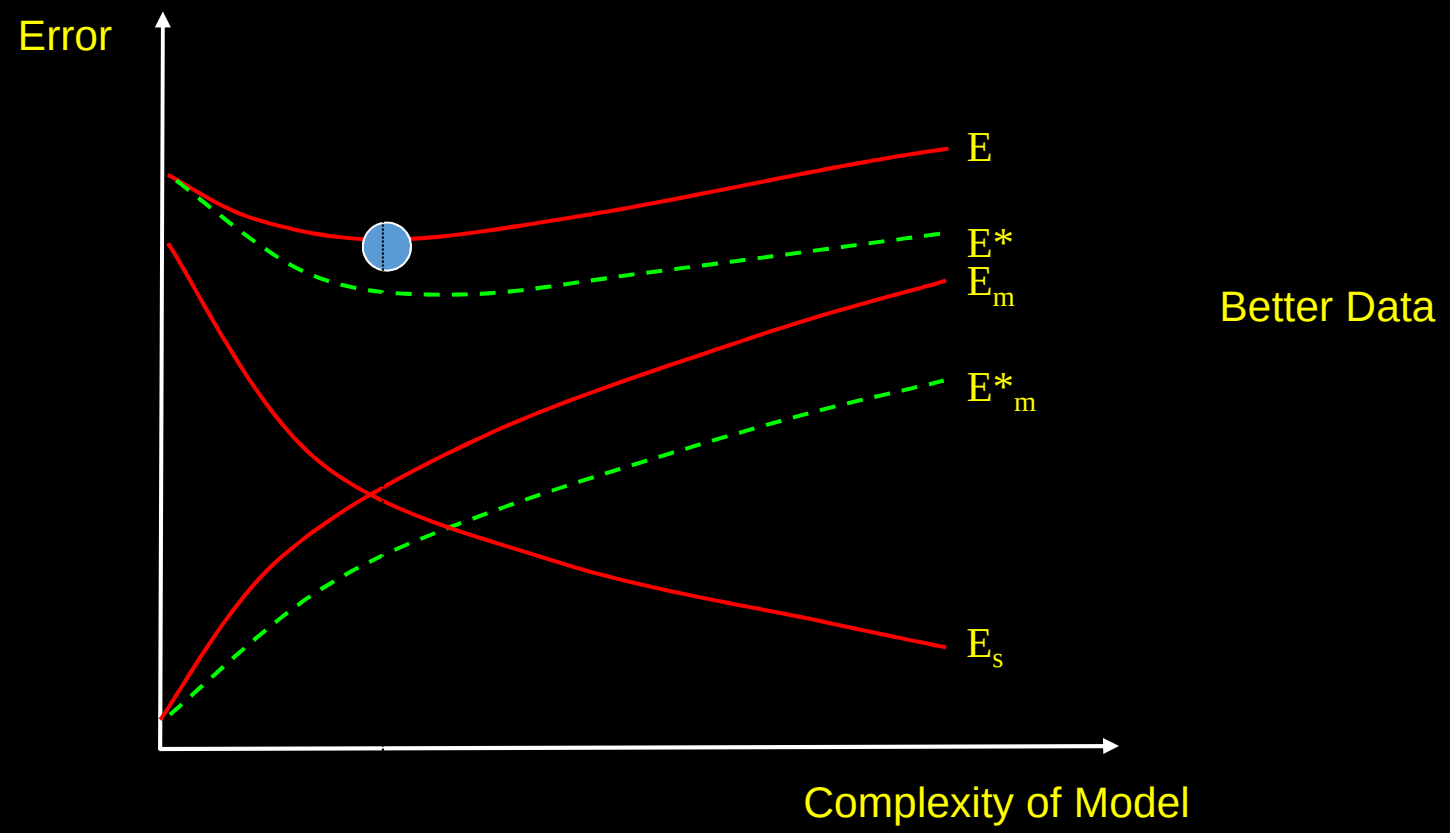
Errors in Models



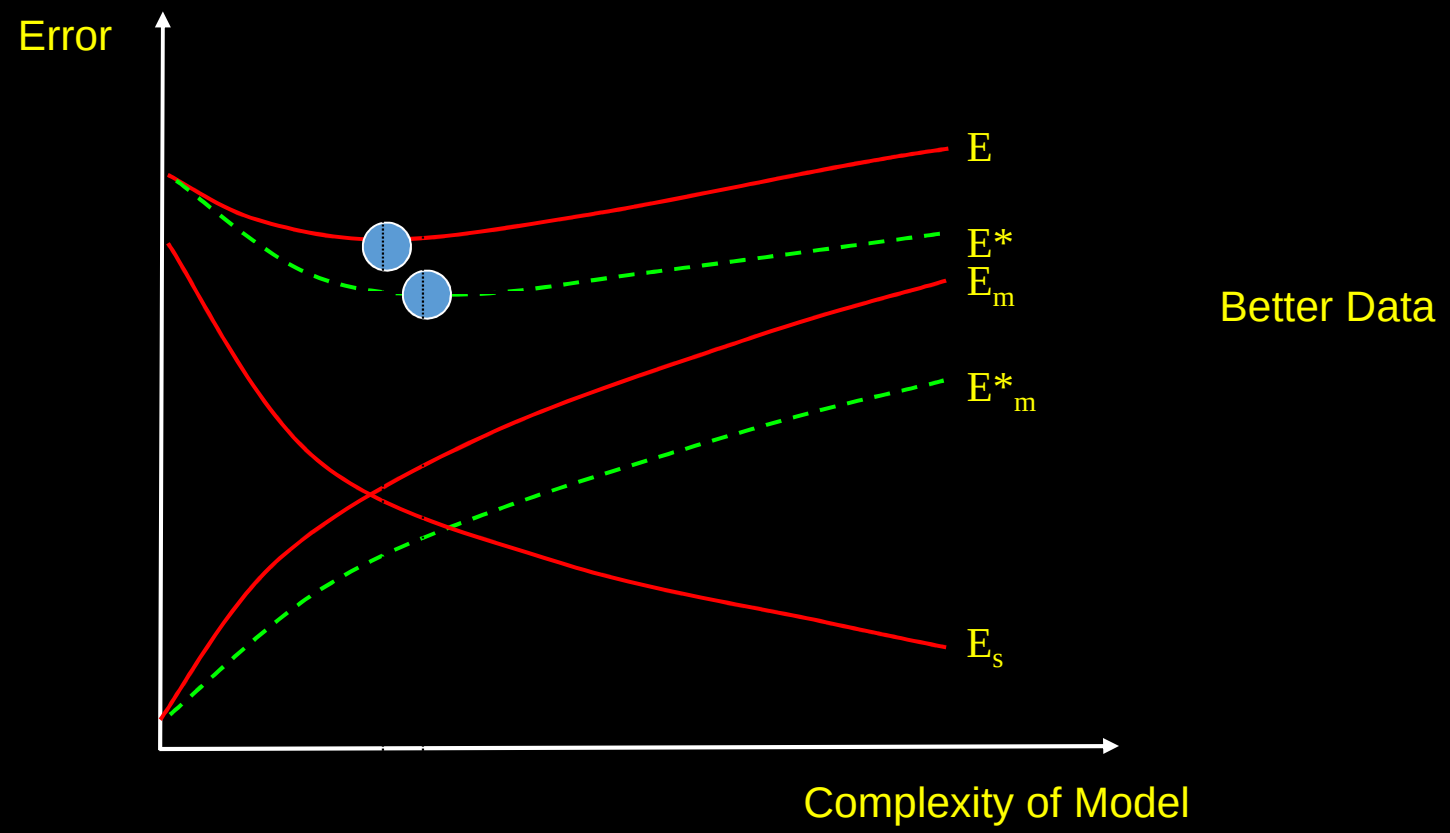
Errors in Models



Errors in Models



Errors in Models



Additional Accumulation of Error

- Errors can also be increased by the manner in which the data is handled.
- Consider starting with data that is accurate to 2% margin.
- Depending how we write our equations we could transform that error term of 2% into 100% or even 300% !

Example

- Assume a population of Sociology students (P_1) = 100 with an error of $\pm$ 2%
- Assume a Population of Anthropology students (P_2) = 102 with an error of $\pm$ 2%
- Further assume that error terms go in the same direction

Example (Cont'd)

- Now look at the errors if I seek an aggregate of the two or the difference between the two values

Example (Cont'd)

$P_1 = 100 \pm 2$ (error term = 2%)

$P_2 = 102 \pm 2$ (error term = 2%)

$P_2 + P_1 = 202 \pm 4$ (error term $\approx$ 2%)

$P_2 - P_1 = 2 \pm 2$ (error term $\approx$ 100%)

Error Terms

If $Z = f(x_1, x_2, \ldots, x_n)$

Then the error term in the function Z will be given by the following equation:

$$e_z^2 = \sum_i fx_i^2 \, e_{x_i}^2 + \sum_i \sum_j f_{x_i} \, f_{x_j} \, e_{x_i} \, e_{x_j} \, r_{ij}$$

Where e_z = error term in Z

$$f_{x_i} = \frac{\partial f}{\partial x_i}$$

e_{x_i} = measurement error in x_i

r_{ij} = correlation between x_i and x_j

Error Terms

If $Z = f(x_1, x_2, \ldots, x_n)$

Then the error term in the function Z will be given by the following equation:

$$e_z^2 = \sum_i f x_i^2 \, e_{x_i}^2 + \sum_i \sum_j f_{x_i} \, f_{x_j} \, e_{x_i} \, e_{x_j} \, r_{ij}$$

Where e_z = error term in Z

$$f_{x_i} = \frac{\partial f}{\partial x_i}$$

e_{x_i} = measurement error in x_i

r_{ij} = correlation between x_i and x_j

Hence…

Seven Rules for Building Models

- Avoid inter-correlated variables

Seven Rules for Building Models

- Avoid inter-correlated variables
- Add whenever possible

Seven Rules for Building Models

- Avoid inter-correlated variables
- Add whenever possible
- If not possible then multiply or divide

Seven Rules for Building Models

- Avoid inter-correlated variables
- Add whenever possible
- If not possible then multiply or divide
- Avoid subtraction and exponentials

Seven Rules for Building Models

- Avoid inter-correlated variables
- Add whenever possible
- If not possible then multiply or divide
- Avoid subtraction and exponentials
- Avoid models that proceed in chains

Seven Rules for Building Models

- Avoid inter-correlated variables
- Add whenever possible
- If not possible then multiply or divide
- Avoid subtraction and exponentials
- Avoid models that proceed in chains
- **Simpler partial models can be more robust than one complex models**

Seven Rules for Building Models

- Avoid inter-correlated variables
- Add whenever possible
- If not possible then multiply or divide
- Avoid subtraction and exponentials
- Avoid models that proceed in chains
- Simpler partial models can be more robust than one complex models
- **Always report predictable error (essential for cases of asymmetrical costs)**

Seven Rules for Building Models

- Avoid inter-correlated variables
- Add whenever possible
- If not possible then multiply or divide
- Avoid subtraction and exponentials
- Avoid models that proceed in chains
- Simpler partial models can be more robust than one complex models
- Always report predictable error (essential for cases of asymmetrical costs)

Envoi

Sound Development Policies Need Thoughtful Social Research Inputs

In a time of remarkable dynamic change, volatility and instability, some new tools may be needed

Natural and social scientists Of the developing world must master the quantitative techniques that have become essential parts of contemporary research…

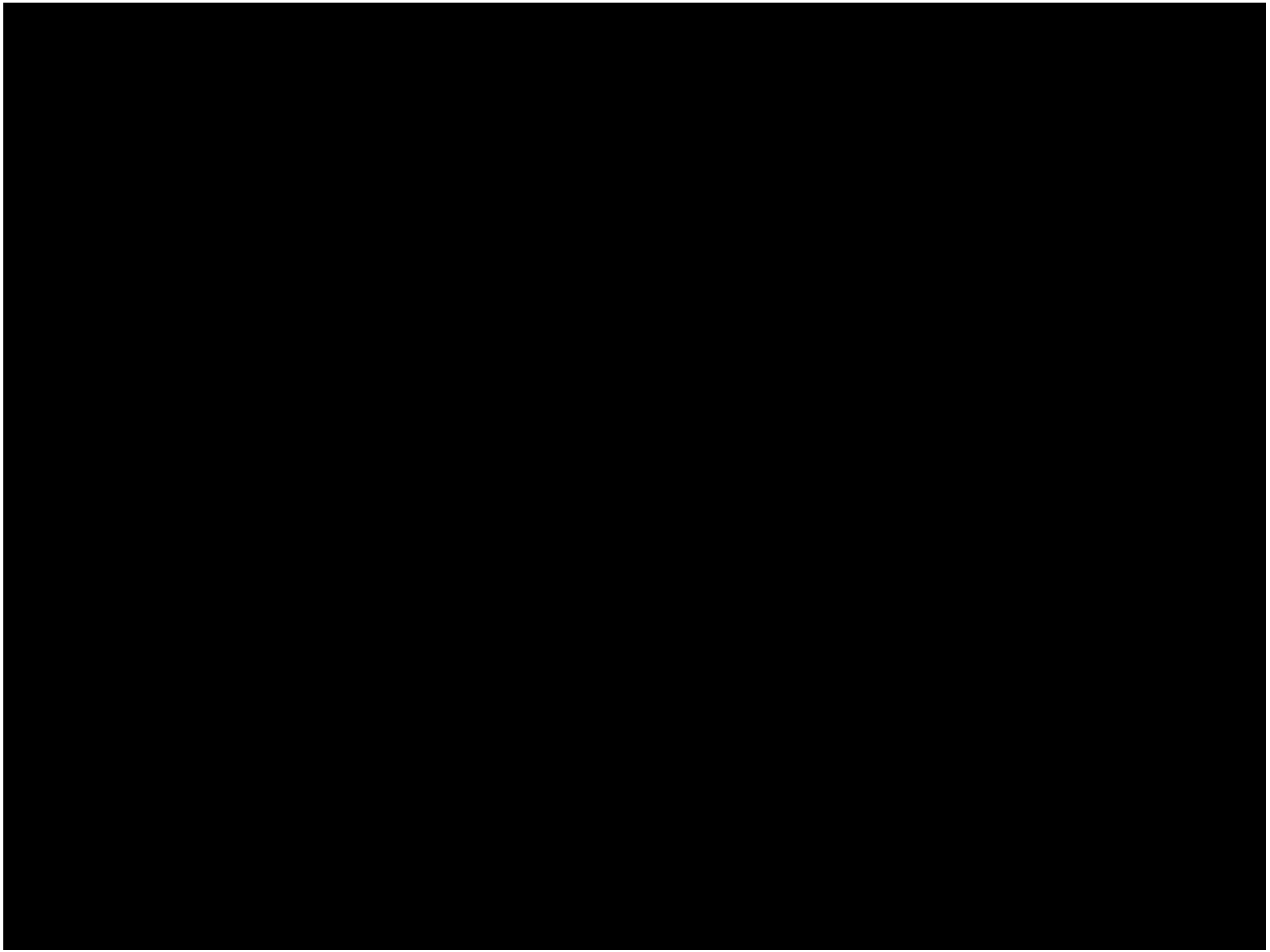
And given the conditions prevailing in much of our education system…

Thus it is not just a matter of publishing learned papers or recognizing past individual achievements…

Thus it is not just a matter of publishing learned papers or recognizing past individual achievements…

it is very much a matter of incubating a revolution!

Thank You



The images used in this presentation are strictly for the educational purpose of this lecture. Any use by anyone for any other purpose should be after consulting the copyright owners of these pictures