Lecture 2

Vladimir Asriyan and John Mondragon

UC Berkeley

September 14, 2011

Theory

Writing a model requires making unrealistic simplifications.

Two inherent questions (from Krugman):

Theory

Writing a model requires making unrealistic simplifications.

Two inherent questions (from Krugman):

- How do we decide what to include?
- How do we decide if it is a good model?

Theory

Writing a model requires making unrealistic simplifications.

Two inherent questions (from Krugman):

- How do we decide what to include?
- How do we decide if it is a good model?

"You make a set of clearly untrue simplifications to get the system down to something you can handle; those simplifications are dictated partly by guesses about what is important, partly by the modeling techniques available. And the end result, if the model is a good one, is an improved insight into why the vastly more complex real system behaves the way it does."
-Paul Krugman

When should we take models seriously?

Saez and Diamond have three requirements:

When should we take models seriously?

Saez and Diamond have three requirements:

- Result should depend on an empirically relevant and important mechanism.

When should we take models seriously?

Saez and Diamond have three requirements:

- Result should depend on an empirically relevant and important mechanism.
- The result should be robust to alternative modelling assumptions.

When should we take models seriously?

Saez and Diamond have three requirements:

- Result should depend on an empirically relevant and important mechanism.
- The result should be robust to alternative modelling assumptions.
- The result should be implementable.

If a model doesn't satisfy these conditions then what is the point?

Student Example

Oportunidades is a program intended to help the poor in Mexico make better health and education decisions. But there is an issue:

Studies have shown that the program does help, but only up to a point. It turns out that when the program gives more money the recipients have **worse** health outcomes!

Question:

Student Example

Oportunidades is a program intended to help the poor in Mexico make better health and education decisions. But there is an issue:

Studies have shown that the program does help, but only up to a point. It turns out that when the program gives more money the recipients have **worse** health outcomes!

Question: Why do health outcomes improve and then decline with the size of the monetary payment?

Student Example

Oportunidades is a program intended to help the poor in Mexico make better health and education decisions. But there is an issue:

Studies have shown that the program does help, but only up to a point. It turns out that when the program gives more money the recipients have **worse** health outcomes!

Question: Why do health outcomes improve and then decline with the size of the monetary payment?
Ideas?

Constructing the model

What are the key elements to any economics model?

Constructing the model

What are the key elements to any economics model?

- Agents: whose behavior are we studying? What do we think guides their behavior?

Constructing the model

What are the key elements to any economics model?

- Agents: whose behavior are we studying? What do we think guides their behavior?
- Environment: what constraints do our agents face? How do they interact with each other? Do we care about many periods, an infinite number, just one?

Agents

Who are the relevant agents?

Agents

Who are the relevant agents? Participants in the program. How do we model them?

Agents

Who are the relevant agents? Participants in the program. How do we model them?

$$\max_{j,h} \alpha \log(j) + \log(h) - \delta\theta c \log(j)$$

- j: junk food consumed.
- h: healthy food consumed.
- $\delta \in [0, 1]$: discount factor (how much does tomorrow matter).
- c: health cost of eating junk food.
- $\theta \in [0, 1]$: "perception" of health costs (1 means I see all the health costs, 0 means I don't see any).

What kinds of assumptions are lurking here? Are we missing any important agents? Can we ask different questions about this program where other agents would be very important?

Environment

What about the environment?

Environment

What about the environment?
Budget constraint! Agents operate in a market economy where they have money (endowments) and make purchases. These two must be equal.

$$\max_{j,h} \alpha \log(j) + \log(h) - \delta\theta c \log(j) \quad \text{s.t.}$$
$$pj + h = M$$

What assumptions have we made here? What are we missing? Is that OK?

What next?

Solve the model (first order conditions):

- Discuss! What forces are at work? What is the role for policy?
- Extend! Are their similar situations you can approach with the same model? Can you address alternative policies?

Economic Inequality

Economic inequality is an important topic:

- "The most profound change in American Society in your lifetime." - Timothy Noah (Slate)
- Did the concentration of wealth contribute to the financial crisis?
- Can we address budget problems with taxing the rich?
- What are the political ramifications of large inequality?

To address any of these questions we need to be able to adequately measure inequality.

Statistics Refresher

Cumulative distribution function (CDF): $F(x)$ Gives the probability that a random variable X has a value less than or equal to x:

$$P(X \leq x) = F(x)$$

Probability density function (PDF): $f(x)$ Function that gives the relative likelihood that a random variable occurs at a certain value x:

$$P(a \leq X \leq b) = \int_a^b f(x)dx$$

Important relationship: $F'(x) = f(x)$

We want to talk about income inequality, so let F describe the income distribution: $F(x)$ gives the fraction of the population with income below x.

Let p be the pth percentile. Then x_p is the level of income such that p percent of the population has income below x_p:

$$F(x_p) = p$$

Measuring Inequality (borrowed from Sen [1997])

Imagine a country with persons $i = 1, \ldots, n$ where y_i is person i's income and $\mu = \sum_1^n y_i/n$ is the average income. The share of person i is given by $x_i = y_i/(n\mu)$.
How do we want to measure inequality?

Range measure:

$$M = \frac{\max_i y_i - \min_i y_i}{\mu} \quad (1)$$

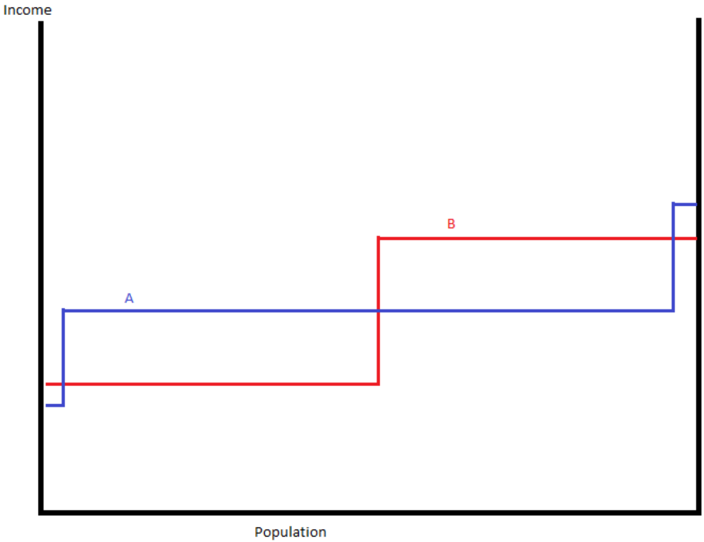
Measuring Inequality (borrowed from Sen [1997])

Imagine a country with persons $i = 1, \dots, n$ where y_i is person i's income and $\mu = \sum_1^n y_i/n$ is the average income. The share of person i is given by $x_i = y_i/(n\mu)$.
How do we want to measure inequality?

Range measure:

$$M = \frac{\max_i y_i - \min_i y_i}{\mu} \quad (1)$$

Problems?


Income
B
A
Population

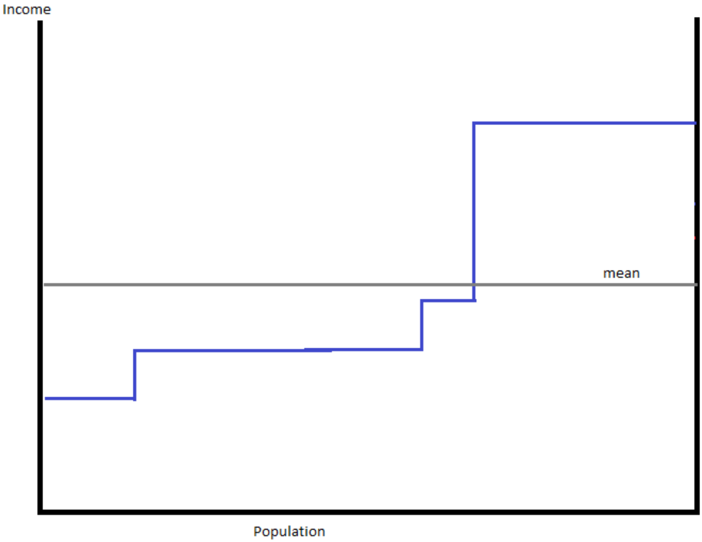
Relative mean deviation:

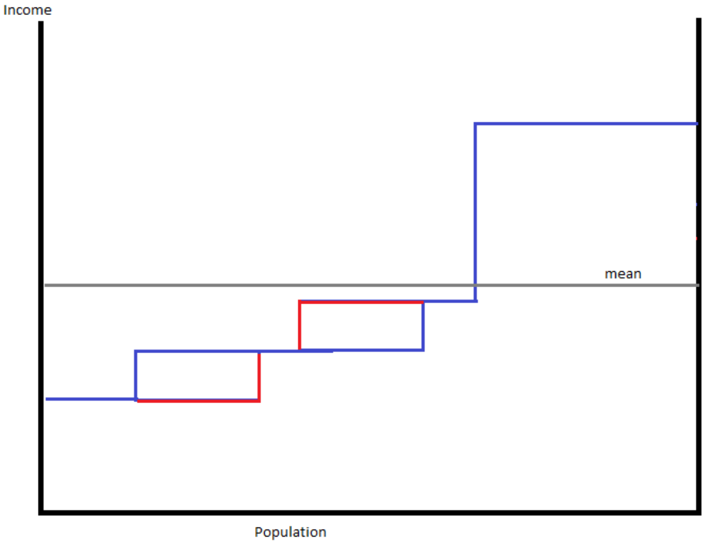
$$M = \frac{\sum_{i=1}^{n} |\mu - y_i|}{n\mu} \tag{2}$$

Relative mean deviation:

$$M = \frac{\sum_{i=1}^{n} |\mu - y_i|}{n\mu} \tag{2}$$

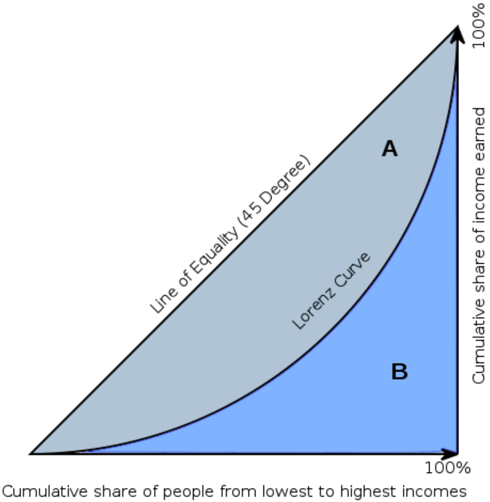
- Takes into account the entire distribution.
- Doesn't capture redistributions on the same side of the mean.


Income
mean
Population


Income
mean
Population

Gini Coefficient

$G = A/(A + B)$



Mathematically, we can define the Lorenz curve as a function $L(F(x))$ where x is a value of random variable X with pdf f and CDF F:

$$L(F(x)) = \frac{\int_{-\infty}^{x} tf(t)dt}{\int_{-\infty}^{\infty} tf(t)dt} \tag{3}$$

The Gini coefficient is defined as

$$G = 1 - 2\int_{0}^{1} L(F(x))dx \tag{4}$$

Think about the extreme cases.

Pareto Interpolation

When actually measuring income we often have data organized into ranges, not individual level data.

TABLE 2
EXAMPLE OF INCOME TAX DATA: UK SUPER-TAX, 1911–12

Income class		Number of persons	Total income assessed
At least	but less than		
£5,000	£10,000	7,767	£52,810,069
£10,000	£15,000	2,055	£24,765,153
£15,000	£20,000	798	£13,742,318
£20,000	£25,000	437	£9,653,890
£25,000	£35,000	387	£11,385,691
£35,000	£45,000	188	£7,464,861
£45,000	£55,000	106	£5,274,658
£55,000	£65,000	56	£3,295,110
£65,000	£75,000	37	£2,590,606
£75,000	£100,000	56	£4,929,787
£100,000	—	66	£12,183,724
Total		11,953	£148,095,867

Source: Annual Report of the Inland Revenue for the Year 1913–14: table 140, p. 155.

We need to make an assumption about distribution of income. A distribution that works for many phenomena is the Pareto distribution:

$$F(x) = 1 - \frac{k}{x}^{\alpha}$$
$$f(x) = \alpha \frac{k^{\alpha}}{x^{1+\alpha}}$$

if $k > 0$ and $\alpha > 0$.

What does this do for us?

- Pick a threshold income y.
- Now take the average of all incomes above y: $y^*(y) = E[Y|Y > y]$
- Now consider the ratio $y^*(y)/y$.
- This ratio is constant for all y!

Specifically:

$$\frac{y^*(y)}{y} = \frac{\alpha}{\alpha - 1} \equiv \beta$$

Not only is this a nice quality (easy to use), but it seems to fit the right tail of the income distribution well.

Saez and his co-authors refer to β as the inverted Pareto coefficient.

- Assume the distribution of the income data is Pareto
- Use the rough data to identify the parameter β (α)
- Use β to recover the hidden data.

Example: If we have estimated that the income distribution looks like a Pareto distribution with $\beta = 2$, then we know that the average income of all individuals with income over 1 million dollars is 2 million dollars.

Figure 1. The Top Decile Income Share in the United States, 1917–2007.

Notes: Income is defined as market income including realized capital gains (excludes government transfers). In 2007, top decile includes all families with annual income above $109,600.

Source: Piketty and Saez (2003), series updated to 2007.

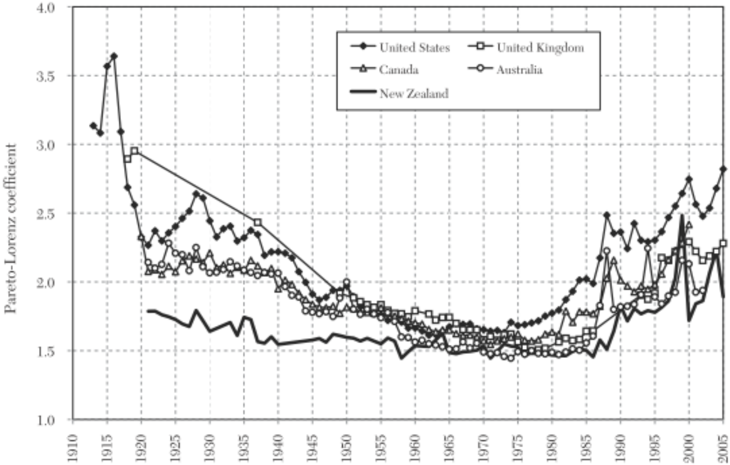


Figure 12. Inverted-Pareto β Coefficients: English-Speaking Countries, 1910–2005

Source: Atkinson and Piketty (2007, 2010).