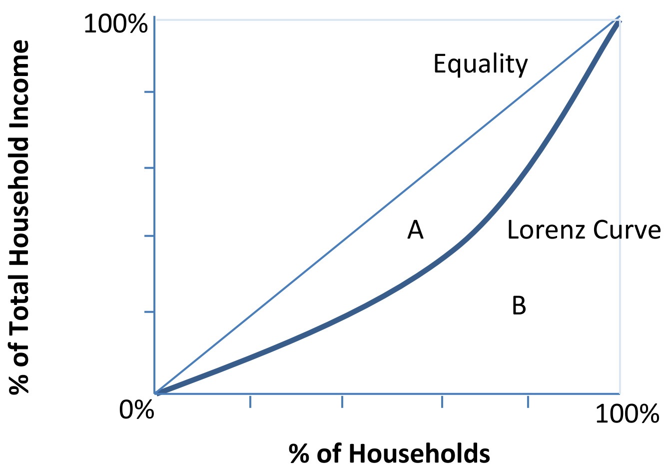
Measuring Distributional Inequalities[1]
Instructional Primer[2]

The Lorenz Curve and Gini Coefficient

The ***Lorenz Curve*** and ***Gini Coefficient*** are tools to help us understand distributional income and/or wealth differences among groups; whether they be nation states or demographically separable groups within a restricted economy.

The ***Lorenz Curve*** compares a group's distributional characteristics to a "line of equality" on an X, Y graph as shown below. The "line of equality" is upward sloping at a 45^0 angle from the point of origin such that it has a slope of 1; that is, for every single unit increase in "rise" there is a corresponding single unit increase in "run" $\left[\frac{1}{1} = 1\right]$. The greater the gap between the ***Lorenz Curve*** and *line of equality*, the greater the distributional inequality.

The ***Gini Coefficient*** compares the area between these curves in such a way as to form a ratio of the area between the ***Lorenz Curve*** and *line of equality* (A) and the area under the *line of equality* (B) (between the *line of equality* and the X axis), such that $\frac{A}{A+B} = Gini\ Coefficient\ \leq 1$. The higher the coefficient, the greater the distributional inequality.



Measuring Inequalities

For simplicity and illustrative purposes we'll look at two groups within an economy (East and West), separated by quintile (five equally represented segments of each group), and consider their relative

[1] This primer is intended to present an abbreviated discussion of the included economic concepts and is not intended to be a full or complete representation of them or the underlying economic foundations from which they are built.

[2] This primer was developed by Rick Haskell, Ph.D. Student, Department of Economics, College of Social and Behavioral Sciences, The University of Utah, Salt Lake City, Utah (2013)

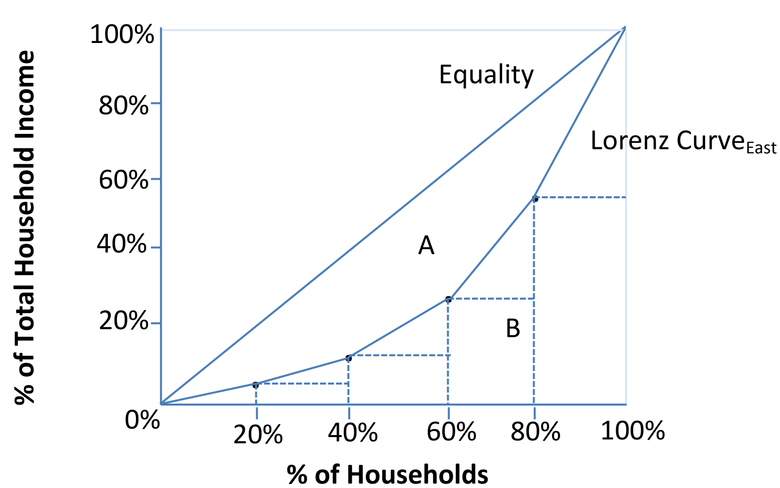
levels of income equality. This requires that we consider the total income of the subject group (Y axis) and the total households of that group (X axis); in each case representing the percentage of total income or percentage of total households represented by each segment within the group. For example: we'll consider the segment of the East group that represents that 20% of households holding lowest percentage of total income and call it the 1st quintile, 20% of households holding next lowest percentage of total income and call it the 2nd quintile, and so forth through the 20% of households holding the highest percentage of household income and call it the 5th, or highest, quintile.

Using the following data we can identify the ***Lorenz Curve*** and ***Gini Coefficients*** for each group:

	East		West	
Quintile	% of Total Household Income	Cumulative % of Total Household Income	% of Total Household Income	Cumulative % of Total Household Income
1st	3.2	3.2	4.6	4.6
2nd	8.5	11.7	10.3	14.9
3rd	15.1	26.8	15.8	30.7
4th	24.7	51.5	23.0	53.7
5th	48.5	100.0	46.3	100.0

When graphing the ***Lorenz Curve*** we are graphing the Cumulative % of Household Income on the X axis as noted in the graph. This allows us to represent a curve that begins at the point of origin and ends at the point where the % of Total Household Income and % of Households each equal 100%. It is important to note that ***Lorenz Curve*** cannot be above the line of equality based on the way that we define the quintile (the 1st quintile will always hold the households with the lowest portion of total household income, etc.).

If we had an equation that represented the shape of the ***Lorenz curve*** we could simply take the integral of the equation to identify these areas (A and B as noted in the following graph), but that's a higher level of math than I'm trying to address here and we don't have that equation available for this example. So we need to think about the "curve" as being a series of straight lines. Let's consider the shape of the ***Lorenz Curve*** for East based on the data above and note the differences between the curve and the *line of equality*:



It's easier to measure the area of B than it is A because B has a constant, horizontal lower bound (bottom) with a vertical right-hand bound (right side) and this makes it easier to begin to envision the area as a series of rectangles and triangles, the areas of which can be calculated and summed to identify the area of B.

Once we know the area of B we can subtract it from the area of A + B to arrive at the area of A. So, it's also helpful to note that the area of A + B is a constant given that it is a right angle triangle that can be calculated as $\frac{base*height}{2} = \frac{100\%*100\%}{2} = \frac{1.0*1.0}{2} = .5 = A + B$. Since we're using quintiles we also know that the base of each rectangle and triangle will also be a constant equal to .20 (each quintile represents 20% of the total).

So let's calculate the ***Gini Coefficient*** for the East given what we know.

The equation for the sum of the rectangles is:

$$(.032)(.20) + (.117)(.20) + (.268)(.20) + (.515)(.20) = sum\ of\ rectanlges \quad (1)$$

We can factor out the .20 and rearrange the equation such that:

$$(.20)(.032 + .117 + .268 + .515) = (.20)(.932) = .1864 \quad (2)$$

$$.1864 = sum\ of\ rectangles \quad (3)$$

Given the data above the equation for the triangles is:

$$\frac{.032*.2}{2} + \frac{.085*.2}{2} + \frac{.151*.2}{2} + \frac{.247*.2}{2} + \frac{.485*.2}{2} = sum\ of\ triangles. \quad (4)$$

We can factor out $\frac{.2}{2} = .1$ such that:

$$(.1)(.032 + .085 + .151 + .247 + .485) = (.1)(1.0) = .10 \quad (5)$$

$$.10 = sum\ of\ triangles \quad (6)$$

Notice that this is also a constant, so we don't even need to calculate it in the future; it is always equal to .10.

Now we can sum the triangles and rectangles such that:

$$.10 + .1864 = .2864 = sum\ of\ triangles + sum\ of\ rectangles = B \quad (7)$$

and we can now subtract B from A+B such that:

$$.5 - .2864 = .2136 = A. \quad (8)$$

Finally we apply the ***Gini Coefficient*** ratio of $\frac{A}{A+B} = Gini\ Coefficient_{East}$ such that:

$$\frac{A}{A+B} = \frac{.2136}{.5} = .4272 = Gini\ Coefficient_{East}. \quad (9)$$

Recall that we expect this to be ≤ 1 and it is.

Let's rely on the constants we identified above to calculate the ***Gini Coefficient$_{West}$*** – this will be a lot easier now.

$$A = .5 - B = .5 - \left(.1 + (.20)(.046 + .149 + .307 + .537)\right) \quad (10)$$

$$A = .5 - (.1 + (.2) * 1.039)) \quad (11)$$

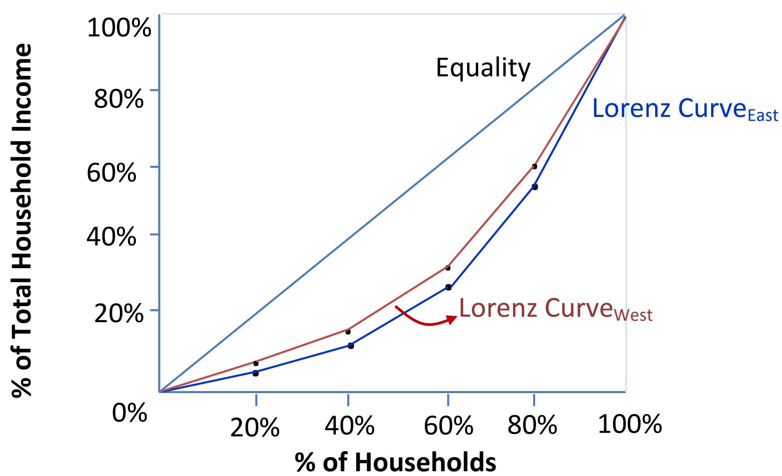
$$A = .5 - (.1 + .2078) = .1922 \quad (12)$$

Recall that this is the same as $.5 - B = sum\ of\ triangles + sum\ of\ rectangles - B = A$

$$\frac{A}{A+B} = \frac{.1922}{.5} = .3844 = Gini\ Coefficient_{West}$$

So… an examination of our two ***Gini Coefficients*** (East and West) reveals that $Gini_{West} \leq Gini_{East}$ (.3844 ≤ .4272) which would suggest that there is greater income inequality in the East than there is in the West.

We can also examine both of the ***Lorenz Curves*** (East and West) on the same graph and we see that the ***Lorenz$_{East}$*** is further from the line of equality that is the ***Lorenz$_{West}$*** supporting the findings of our ***Gini Coefficients*** for these two groups.



Other Distribution Ratios

It is common to specify other ratios in respect to distributional equalities, such as the ratio of the highest income group to the lowest (highest: lowest or $\frac{highest}{lowest}$). These take various forms, but it is relatively standard to represent the lower income (wealth) group in the denominator and the higher income (wealth) group in the numerator. Using the East/West data above, we might want to compare the income ratio of the highest quintile to the lowest quintile for each group. Each ratio tells us something of import regarding the distributional equality of that group, but a comparison of the ratios for each group is also meaningful. For example, a ratio comparing the top quintile to the bottom for East lets us see how disparate the income distribution of that economy is. A comparison of that ratio for the East versus the same ratio for the West tells us the relative distributional income equality of the two groups.

	East	**West**
Top quintile to lowest quintile	$\frac{48.5}{3.2} = 15.15$	$\frac{46.3}{4.6} = 10.06$
Top 40% to lowest 40%	$\frac{24.7 + 48.5}{3.2 + 8.5} = \frac{73.2}{11.7} = 6.25$	$\frac{23.0 + 46.3}{4.6 + 10.3} = \frac{69.3}{14.9} = 4.65$

The 1st row interpretation is that those in the upper most income quintile in the East have incomes roughly 15 times greater than those in the lowest quintile, and the upper most quintile in the West have incomes roughly 10 times greater than those in the lower quintile. This supports the previous finding (as represented by the ***Gini Coefficients*** and ***Lorenz Curves***) that there is greater income inequality in the East than in the West.

The 2nd row interpretation is interesting in that it makes the two groups appear to have somewhat similar distributional properties, when in fact, they do not. So… it is always important to be clear as to what you're comparing and to also be aware if it is giving you the most information possible.

When comparing segments of two or more groups it's important to always compare like-for-like, otherwise the comparison is of little value. For example, we wouldn't gain any meaningful data by comparing the upper most quintile against the lowest for one group and the upper 50% against the lowest 50% of another.

Finally, we've considered percentage income distribution in these examples, but it is also common to compare nominal ratios or real ratios. The ratio form you employ will largely be a function of the data at your disposal or, in the case of the ***Gini Coefficient*** and ***Lorenz Curve***, the required form for the model.

Lorenz Curve and Gini Coefficient Problem Set

Use the information provided for groups X and Y to construct answers for parts a and b.

a. Construct one chart showing the Lorenz Curve for each group using quintile segregation. Which group appears to have the greater level of income inequality?

Group X		Group Y	
Becca	42000	John	15000
Bill	20000	Beth	30000
Charles	31000	Gina	30000
Kasey	24000	Joe	74000
Nina	34000	Julie	22000
Raul	37000	Leslie	60000
Rose	29000	Mark	38000
Tom	35000	Rick	25000
Will	60000	Sheri	18000
Yukiko	32000	Steve	32000

b. Calculate the Gini Coefficient for each of the two groups.