

Study on the Methods of Income Inequality Measure

Guilan Huang

*Nanjing Xiaozhuang University, Jiangsu, China
sandybhgl@163.com*

Keywords: Gini coefficient, distribution gap, Lorenz curve, Origin by computer

Abstract: The measure of income inequality is the basis of weighing it and key elements to make comparisons regionally and internationally by Origin of computer. Of all the different methods, the Lorenz Curve and Gini Coefficient, Theil's Entropy Measure and Toxic Index Method are the most used ones. However, during the process of using them to carry on all the comparisons, the applicable standards are hard to meet. Instead of putting one single index into effect, only when using them in combination can we achieve better results.

1. Introduction

Income inequality is the existence number features of the income distribution gap. The study on income inequality methodology is the basis of measuring of the extent of income inequality. Scholars have brought forth many practical and commonly used methods, such as Lorenz Curve, Gini coefficient, Toxic Index Method, Kuznets Index, the "Inverted U" Curve, Halving Method, Engel Coefficient Method, entropy index method, etc. Allowing for Kuznets Index and the Inverted U Curve incline to measure the income inequality between different regions rather than groups, Engels coefficient is used to calculate the food accounts for the proportion of total expenditure, which is usually used as one of the income gap between the international comparison of indicators. Whereas both developing and developed countries, The fund of food expenditure to maintain the livelihood of a family have a rigidity character. It's also probably incredible due to the differences between household purchasing power and choices in the quality of food. Lorenz Curve and Gini Coefficient's method and numerical can more objectively reflect a country's overall income distribution gap and the internal distribution of income difference among residents. In the process of its development according to the survey data availability and data differences, more operable method are developed to carry on calculations. It is the gini coefficient of horizontal comparison, exploitativeness of practical, easy access to data that make it become the most common measure in the world. Because researches on personal income tax redistribution effect of quantitative are based on Gini coefficient, we regard it as the basis of income inequality research tools. Rather, it is quite sensitive to changes of the wages of moderate income group and dull to of the income of low-income and high-income. In the study of income inequality, we will measure all levels of internal and income inequality with Theil's Entropy Measure and Toxic Index Method being supplement.

2. The Lorenz Curve

Lorenz (1905), when studying the fairness of the income distribution of production factors, categorized the cumulative percentage of the population of a country in terms of wealth (income) from low to high, and the distribution of the population's corresponding wealth to the total social wealth as a variable. On the horizontal and vertical axes of a square, the Lorenz curve of the country's income distribution for the year is plotted.

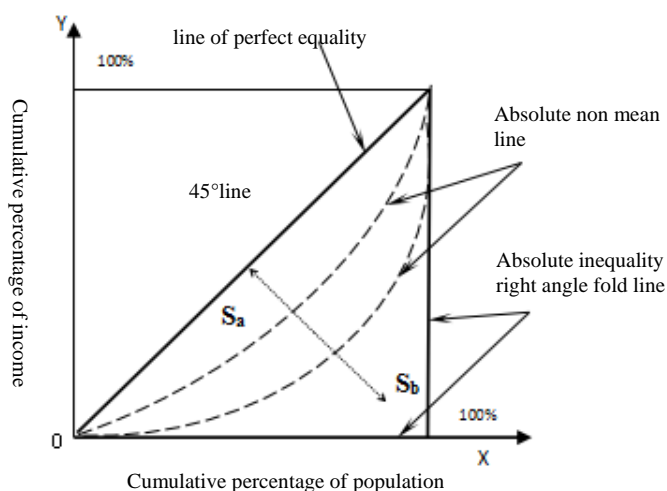


Figure 1 Lorenz curve.

2.1 The characteristics of the Lorenz curve

As shown in Figure 1, the Lorenz curve starts from point (0,0) and ends at point (1,1); when the Lorenz curve is an absolute average line of 45 degrees, it indicates that the income distribution in the country is absolutely fair, and the per capita Absolute average income levels; when the Lorenz curve is a vertical straight-line composite line with a horizontal axis and a population of 100%, it shows that the income distribution in the country is absolutely unfair, and the per capita income level is absolutely uneven. The 45° absolute average and the absolute-unequal right-angle poly line are two extreme income distribution states. Normally, the Lorenz curve lies between the 45° line and the right-angled poly line, which is an absolute non-average line. When the Lorenz curve approaches the 45° line, it indicates that the income distribution tends to be absolutely fair. When the Lorenz curve approaches the right-angled poly line, it indicates that the income distribution tends to be absolutely unfair.

2.2 The fitting method of Lorenz curve equation

Although the Lorenz curve more intuitively describes income inequality, there is no uniform measurement (measurement) equation. An accurate Lorenz curve can improve the accuracy of the study. In general, there are three methods for fitting the Lorenz curve: geometric calculation, indirect fitting, and curve fitting. Geometric calculation method is based on statistical data in accordance with the block data for block fitting, the accuracy is poor; indirect fitting method is based on the premise of the known income distribution probability density function, to obtain the Lorenz curve equation; curve fitting method is The continuous Lorenz curve is fitted by the grouping data. Although there are errors, the accuracy is higher than the geometric calculation method and indirect fitting method. It is an ideal method.

3. The Gini Coefficient

Gini(1912) According to the Lorenz curve, the concept of the Gini Coefficient and the specific calculation method were proposed. The degree of inequality of income distribution was determined in the form of quantity, which compensated for the quantitative loss of the Lorenz curve to income inequality. The Gini coefficient compares the area of unequal income distribution in the Lorenz curve with the area enclosed by absolute equality and absolute inequality to measure the degree of inequality in income distribution. In this way, the Gini coefficient is calculated as

$$Gini = \frac{S_a}{S_a + S_b}$$

Because of $S_a + S_b = \frac{1}{2}$, $Gini = \frac{S_a}{S_a + S_b} = 2S_a$, among $0 \leq S_a \leq \frac{1}{2}$, When $S_a = 0$, there were $S_b = \frac{1}{2}$, indicating that the income was absolutely fair, that is, the Lorenz curve was a 45° line.

When $0 < S_a < \frac{1}{2}$, $0 < Gini < 1$, it showed that the income gap was unfair, that is, the Lorenz curve was an absolute non-average line between the 45° line and the straight line.

When $S_a = \frac{1}{2}$, there were $S_b = 0$, $Gini = 1$, indicating that the income was absolutely unfair, that is, the Lorenz curve was an absolutely uneven rectangular line of income.

$$Gini = \frac{S_a}{S_a + S_b}$$

Under the conditions known by the Lorenz curve, the Gini coefficient is

$$\begin{aligned} Gini &= \frac{\int_0^1 X dX - \int_0^1 Y(X) dX}{\int_0^1 X dX} = 1 - \frac{\int_0^1 Y(X) dX}{\int_0^1 X dX} \\ &= 1 - \frac{1}{2} \int_0^1 Y(X) dX \end{aligned}$$

Among them, $\int_0^1 X dX$ is the area enclosed by the 45° line and the absolute uneven rectangular line, that is $S_a + S_b$, $\int_0^1 Y dX$ is the area S_b enclosed by the absolute unbalanced line and the absolute unequal right-angled folding line. In the unknown condition of the Lorenz curve, (1912) is

$$Gini = \frac{\sum_{i=1}^l \sum_{j=1}^l |Y_i - Y_j|}{2l^2 \gamma}$$

Among them, Y_i and Y_j are the incomes of NO.i and NO.j populations, $Y_i - Y_j$ is the income difference between the NO. i and NO. j populations, h is the income expectation value, and the Gini coefficient indicates the extent to which the income gap deviates from the expected income value.

Lorenz (1912) is called the Gini coefficient by the latter. Although this calculation method is relatively intuitive, it requires a large amount of income data and auxiliary calculations. In practice, sample survey data are mostly used. A smaller number of groups will greatly reduce the accuracy of the Gini coefficient calculation. To this end, domestic and foreign scholars have summed up the more practical direct methods, area methods and equidistant methods on the basis of previous studies. Among them, the area method has an upper trapezoidal area method and a lower trapezoidal area method.

3.1 Upper trapezoid area method

As show in Figure 2, the upper trapezoidal area method obtains S_a to calculate the Gini coefficient. Suppose population and wealth or income data are divided into n groups according to the population size $P_i (P_i \neq P_j)$, the total income of Group i corresponding to P_i is Y_i , the proportion of NO. i population in the total population is $p_i = P_i / \sum_{i=1}^n P_i$, and the population of Group I accounts for the total income. If the proportion is $y_i = Y_i / \sum_{i=1}^n Y_i$, then the degree of income distribution inequality S_a is equal to the sum of a number of rectangular areas and a number of triangular areas minus the area of the right triangle above the 45 degree line.

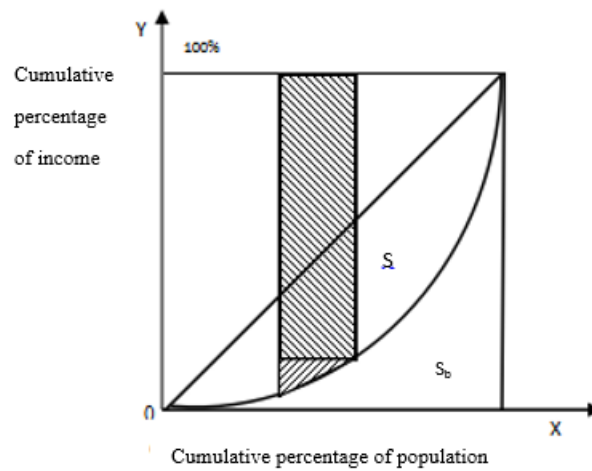


Figure 2: Lorenz curve.

According to the upper trapezoid area method, S_a is

$$\begin{aligned}
 S_a &= \frac{1}{2} p_1 y_1 + p_1 (1 - y_1) + \frac{1}{2} (p_2 - p_1) (y_2 - y_1) \\
 &+ (p_2 - p_1) (1 - y_2) + \frac{1}{2} (p_3 - p_2) (y_3 - y_2) \\
 &+ (p_3 - p_2) (1 - y_3) + \dots + \frac{1}{2} (p_n - p_{n-1}) (y_n - y_{n-1}) \\
 &+ (p_n - p_{n-1}) (1 - y_n) - \frac{1}{2} \\
 &= \frac{1}{2} \sum_{i=1}^n (p_i y_{i+1} - p_{i+1} y_i) + p_n - \frac{1}{2} p_n y_n - \frac{1}{2}
 \end{aligned}$$

Because $Gini = 2S_a$, $p_n = y_n = 1$, Gini coefficient is

$$\begin{aligned}
 Gini &= 2S_a = \sum_{i=1}^n (p_i y_{i+1} - p_{i+1} y_i) + 2p_n - p_n y_n - 1 \\
 &= \sum_{i=1}^n (p_i y_{i+1} - p_{i+1} y_i)
 \end{aligned}$$

If the population is divided into k groups and $P_1 = P_2 = P_3 = \dots = P_k$, and then the Gini coefficient is

$$Gini = \frac{2}{k}(y_1 + 2y_2 + \dots + ky_k) - \frac{k+1}{k}$$

$$= \frac{2}{k} \sum_{i=1}^k iy_i - \frac{k+1}{k}$$

3.2 Lower trapezoid area method

The lower trapezoidal area method is to calculate S_b , and then according to $S_b + S_a = \frac{1}{2}$, draw the Gini coefficient, where S_b is the sum of several trapezoidal areas.

$$S_b = \frac{1}{2}p_1y_1 + \frac{1}{2}(p_2 - p_1)(y_1 + y_2) + \frac{1}{2}(p_3 - p_2)(y_2 + y_3)$$

$$+ \dots + \frac{1}{2}(p_n - p_{n-1})(y_{n-1} + y_n)$$

$$= \frac{1}{2}p_1y_1 + \frac{1}{2} \sum_{i=1}^{n-1} (p_n - p_{n-1})(y_{n-1} + y_n)$$

$$Gini = 2S_a = 2\left(\frac{1}{2} - S_b\right) = 1 - 2S_b$$

$$= 1 - p_1y_1 + \sum_{i=1}^{n-1} (p_n - p_{n-1})(y_{n-1} + y_n)$$

3.3 Direct method

According to the Gini formula of (1912), a more practical direct method was deduced.

$$Gini = 1 - \sum_{i=1}^n p_i \times \left(2 \sum_{k=1}^i (y_k - y_i) \right)$$

In addition to the trapezoidal area method, the lower trapezoidal area method, the rectangular area method, and the fitted curve method, the calculation methods of the Gini coefficient are different. Since the rectangular area method is not used in this paper, it will not be introduced. The fitting curve method is based on the calculation of the Gini coefficient based on the Lorenz-fitting curve.

4. Taylor's Entropy Index Method

Theil(1967) introduced the entropy in information theory into the quantitative measurement of inequality of income distribution and deduced the Theil entropy index (Theil index). The Taylor's entropy index measures the degree of contribution of income inequality to total income inequality among different income groups, individuals or different groups. The formula is:

$$Theil = \sum_{i=1}^n y_i \log\left(\frac{y_i}{p_i}\right)$$

According to the different stratum's of income groups, the Taylor's entropy index can be divided into the Theil's entropy index and the Theil's entropy index. The Theil's entropy index and the Theil's entropy index are more sensitive to the changes in the income of the high-income class and the low-income class, respectively. Its formula is:

$$T = \sum_{i=1}^n y_i \log ny_i \quad L = \frac{1}{n} \sum_{i=1}^n \log \frac{\bar{Y}}{Y_i}$$

Among them, \bar{Y} is the average of total income, that is
$$\bar{Y} = \frac{\sum_{i=1}^n Y_i}{n}$$

5. Bad Income Index Method

The Bad Income Index, also known as the Ox-ma Index (Income unhealthy index), is based on the equidistant method. It uses the ratio of the highest 20% of the income level to the lowest 20% of the population as an indicator of income distribution. for

$$IUI = \frac{M_5}{M_1}$$

Among them, M_1 is the average income of the highest income group, M_5 is the average income of the lowest income group.

When the average income of the lowest income group is equal to the highest average income group, the bad income index is 1, indicating that the income distribution is absolutely fair; the greater the bad income index, the more unfair the income.

Acknowledgements

This work is supported by the Philosophy and Social Science Funds for Jiangsu (Grant No. 2017SJB0427), the Fundamental Research Funds for Nanjing Xiaozhuang University Universities (Grant No. 2017NXY32).

References

- [1] Anthony, A.B. (1970) *On the measurement of inequality*, *Journal of Economic Theory*, 1970(2).
- [2] Kakwani, N. (1970) *Measurement of Tax Progressivity: An International Comparison*. *Economic Journal*, (87):71-80.
- [3] Sung, M.J., Park, K. (2012) *Effects of taxes and benefits on income distribution in Korea*. *Review of income and wealth*, 57(2):345-363.
- [4] Thomas, P., Emmanuel, S. (2007) *How progressive is the U.S. Federal Tax System? A Historical and international perspective*. *the Journal of Economic Perspectives*, 21(1):3-24.
- [5] Urban, I, Lambert, P.J. (2005) *Redistribution, Horizontal Inequity and Reranking: How to Measure Them Properly*. *Public Finance Review*, 36(5):563- 587.
- [6] Wagstaff, A. (2005) *Decomposing Changes in income inequality into vertical and horizontal redistribution and reranking, with applications to China and Vietnam*, *World Bank Policy Research Working Paper*NO.3559.