

## [CN] Xueqiu Bai

### **Contact Information:**

Address: School of Marxism, Peking University, Beijing 100871, China

Telephone:

Email Address: [baixq@pku.edu.cn](mailto:baixq@pku.edu.cn)

Dr. Xue-Qiu Bai is an associate Professor at School of Marxism, Peking University , and the Director of Economics Division in this School. She is also an authors of 40 papers in the peer-refereed journal and a monograph, Economic Development Of China : Overall arranging of urban and rural areas. Her majority of researches is evelopment economics and the rural problem in Chinese economics .

## **Gini Coefficient Analysis on Urban-rural Income Differences**

[CN] Xueqiu Bai

**Abstract:** Some novel theoretical investigations are made for the “one country, two measurements” viewpoint and its controversy with respect to the inequality of the income distribution of Chinese urban-rural dual economy. We build an analytic model based on the Gini coefficient for the dual structure, and then introduce the generalized Lorenz Curves. From this angle of view, the intrinsic relationships are revealed between the urban-rural difference of income and the inequality of income distribution in whole economic society. The Simpson's Paradox is found to exist in Gini coefficient analysis. For example, even though the Gini coefficients of each party vanish, i.e., each of urban-rural parties has perfect equality, there still exists non-zero Gini coefficient for the whole economic society. The whole Gini coefficient is determined propositionally by the urban-rural difference of income. Only the Gini coefficient of a party can not reflect exactly the situation of inequality in the urban-rural dual economy. Therefore, to promote social justice and social harmony we should balance urban and rural development and narrow the income gap between urban and rural residents.

**Key words:** income distribution Gini coefficient inequality urban and rural parts

Economic growth can not be equated with economic development simply, and also the overall development of a country can't be measured just by GDP. In the point of view of the development economics, the development goal of a social is not only to pursue the growth of economy, but also to promote the progress of the society, especially to reduce and eradicate the poverty. It's not a real progress for a society if the raise of the per capita income accompanied by a widening gap between the rich and the poor. For this reason, the distribution of income and wealth among different individuals and groups becomes an extremely important indicator to measure the development of economy and the progress of society. In the theory of economics, it relates to the problem of what kind of functions of social welfare can be suitable to measure the progress of social and which type of indicators can be used to describe the equity of social. Since economic reform and opening up, the economy keeps on growing by an annual average rate of more than 9.5 percent in china, but the income gap is widening with the rapidest rate in the world at the same time. Therefore, the social equity issue about income distribution indeed becomes a core problem in the study of how to build a harmonious society in china. In fact, the inequality of income distribution in a certain period will reduce the possibility of the overall economy to make a further growth. Moreover, a greater inequality of economy can even lead to the social turbulence and completely impede the development of economy. Thus, not only must we solve the problem of inequality from the perspective of the efficiency of long-term economic growth, but also we should attach importance to the solving of unfair income distribution, from the concept of people-centered scientific development, as well as the

humanistic care of social equity, in order to build a harmonious society in which the fruits of development can be equally shared by all the people.

Since the issue of economic inequality is so important, it would be necessary to explore the ways to improve the situation of inequality. Therefore, we need to make it clear firstly: What is the economic inequality and what indicator can be used to measure it? Does it have a recognized standard to evaluate the degree of economic inequality? It must be pointed out that, although the modern economics has not completely resolved this issue, that is, has not yet found a universal indicator to measure inequality, but it still has established some of the generally accepted means to measure the economic inequality. And one of them is the Gini coefficient.

The main motivation of studying the calculating Gini coefficient of urban and rural parts in this paper is to address to the view of “one country, two measures” and its arguments with respect to the inequality of the income distribution, and attempt to provide a rigorous economics’ thinking theoretically. According to the World Bank, China's Gini coefficient was 0.16 before the economic reform and opening up, and reached 0.458 in 2003, which had exceeded the internationally recognized warning line 0.4, then, it kept on expanding and was more than 0.465 in 2004, in 2005 it reached 0.47. Obviously China's Gini coefficient has already exceeded reasonable limits,<sup>1</sup> and has the trend to expand continually. In this regard, some scholars believe that China is a typical country with dual economic structure. So we can't analysis the problem of income distribution in the general standard of Gini coefficient. In their view, as the rural population accounts for the majority, we should adopt the means of “one country, two measures”, calculate respectively in the urban and rural, as the method to measure whether the income distribution is fair or not. In this paper, I will discuss this issue based on the study of the measure of inequality of the

---

<sup>1</sup> It must be noted that the announced Gini coefficient didn't count in the unreasonable, non-standardized, abnormal and illegal income which trigger people's dissatisfaction. Otherwise, the Gini coefficient will be more than 20% larger than the number published at present.

income distribution between the urban and rural in a theoretical point of view. To say the least, even take the means of “one country, two measures” to calculate separately, even the result is that the income distribution is relatively equal between the urban and rural parts<sup>1</sup> and there is a relatively temporary stability, we can’t ignore the truth that the Gini coefficient of the whole community has overrun the warning line. Judging from the economic principle, the method of “one country, two measures”, which is a logical paradox, divides up the whole country's economy in fact. If look at issues on the basis of such perspective, it not only go against the construction of a harmonious society which the fruits of development be shared by all the people, but also deviate from the idea of human-oriented scientific development.

So, I disagree with the method of “one country, two measures”. Here I will start with the general theoretical framework and give a clear definition to the Gini coefficient of the urban and rural parts on the basis of the geometric description of the usual Gini coefficient. And then, explore the generalized Lorenz curve which can accurately reflect the income and the characteristic of equity of the urban and rural through the function of the income of different individuals rearranged by the two parts. Then, the core conclusion can be proved according to the Lorenz curve based on the urban and rural parts, that is, the Gini coefficient of the whole community has a plus amount in addition to the amount weighted according to income ratio in each part. The existence of this extra part means that even the Gini coefficients of each part is zero (absolutely equal within the two parts), the Gini coefficient of the whole community can’t be zero. This is precisely because of the existence of the plus can the inequality between the urban and rural been presented. Finally, I

---

<sup>1</sup> According to the data of National Bureau of Statistics, the Gini coefficient within the urban was 0.23 in 1988 and only increased to 0.319 in 2002, still far from the warning line of 0.4. The Gini coefficient within the rural was larger than the urban, 0.303 in 1988 and increased to 0.366 in 2002, hasn’t reached the level of 0.4 similarly.

discussed some ways and means of narrowing the income gap between the urban and rural.

An relatively broad definition of economic inequality measure is the Lorenz curve and the corresponding Gini coefficient (Ray, 1997). Although the Gini coefficient has various of flaws and fetching errors, as a general international target, the effectiveness of measuring social income distribution differences was recognized by the majority of countries for the past century. Chinese scholars have never questioned on the "Gini coefficient" theory and methods as well. Institute "revision" is just the use of it, such as implementations of "one country, two measures" or "the idea of the urban and rural area division" and so on. Moreover, no better alternative measurement has been found yet still now. Therefore, in this paper we still embark to the urban-rural income distribution conditions research from the Gini coefficient concept and give out the precise definition and detailed presentation of Gini coefficient of the two urban-rural departments so as to reveal the Gini coefficient relationship between the whole society and the urban-rural departments.

Established an economy with  $N$  independent economic entities, represented by  $1, 2, 3, \dots$ , and the income of each is  $y_1, y_2, y_3, \dots, y_N$ . Without losing generality, they are arranged from small to large. Gini coefficient could be defined by the income proportion function  $f(x) = \sum_{k=1}^{xN} y_k / Y$  and the corresponding Lorenz curve, where  $x = n / N$ , represents the proposition of  $n$  personalities to the total population  $N$  and  $Y = \sum_{k=1}^N y_k$  represents the total income of society.

$$G = 1 - 2 \int_0^1 f(x) dx \quad (1)$$

This integral expression form of Gini coefficient is only to indicate conveniently, and for its concrete inferential reasoning see the appendix. Obviously, when  $f(x) = x$  ( $0 \leq x < 1$ ),  $G = 0$ , which means complete average and the income of every person is the same; when  $f(1) = 1$ ,  $G = 1$ , which means complete imparity and those with highest income occupied the whole income. Certainly, this is two extreme cases will exist theoretically but not in reality.

In the above traditional analysis on the Gini coefficient and the Lorenz curve, the fairness of the income distribution is described without distinguishing the economic entity. Formally this is requested by anonymous principle of the definition of the inequality measurement index. That is the individual is ranked only by income, rather than who is the individual, or which category the individual belongs to. Although such treatment can reflect the overall socio-economic inequalities, it could not reflect the inequalities between different groups, because two departments with significant differences in the revenue share

may have the same Gini coefficient.

In this article we assume that the total rural income is  $Y_1$ , and there are  $M$  economic entities with respective income of  $(y_1, y_2, y_3 \dots y_M)$ ; the total urban income is  $Y_2$ , and there are  $N$  economic entities with respective income of  $(y_{M+1}, y_{M+2}, y_{M+3} \dots y_{M+N})$ . The rural and urban income is sorted from the poorest to the richest respectively. The urban-rural department income ratio function  $f(x) = (\sum_{k=1}^{(M+N)x} y_k) / Y$  is directly defined by the total population  $M + N$  and the total income  $Y = Y_1 + Y_2$ . Here, the income of each sector is

$$Y_1 = \sum_{k=1}^M y_k \quad Y_2 = \sum_{k=M+1}^{M+N} y_k = \sum_{k=1}^N y_{k+M}$$

respectively. Therefore, the urban and rural department income ratio function can be defined respectively as a function of the proportion of the population. When  $0 < x < M / (M + N)$

$$f(x) \equiv f_1(x) = \frac{1}{Y} \sum_{k=1}^{x(M+N)} y_k \quad (2)$$

When  $M / (M + N) \leq x < 1$

$$f(x) \equiv f_2(x) = \frac{1}{Y} \sum_{k=1}^{x(M+N)} y_k = \frac{Y_1}{Y} + \sum_{k=M+1}^{x(M+N)} \frac{y_k}{Y} \quad (3)$$

Obviously,  $f(x=1) = 1$ . The economic significance of these expressions above is very obvious.

The promoted Lorenz curve  $OBKCA$  is showed in Figure 1: there is a non-smooth transition point  $K$  in curve  $Y_1$ . If the income within the urban and the rural sector is the same respectively, the arc  $OBK$  and arc  $KCA$  in the Lorenz curve will change to the straight line  $OK$   $KA$  respectively. This is a case worthy of consideration. If the society is completely equal, the whole Lorenz curve  $OBKCA$  will change into a straight line  $OA$  with an angle of  $45^\circ$ , which shares the intersection point  $F$  with line  $EP$ . From the figure above, it could be revealed that the population proportion of the first department to the total population is  $x = M / (M + N)$ , and the corresponding income ratio is  $f(x = M / (M + N)) = Y_1 / Y$ . When the income ratio is not the equal with the population proportion, the curve  $OPFR'$  in the figure will be not a square.

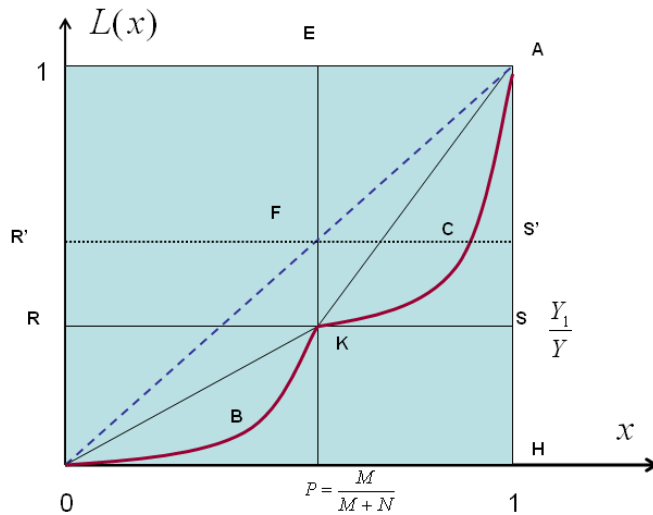


Figure 1: the Lorenz curve on dual economy structure: the urban and rural population arranged cumulative percentage of income from low to high, the rural first and urban after.

The general structure of income distribution function (the Lorenz curve) of a dual economy structure is analyzed above. Now, the society fairness and the inherent relationships between inherent fairness could be analyzed on the basis of its properties. Therefore, the whole community's Gini coefficient is defined with the standard approach in this paper.

$$G = 1 - \frac{2}{Y} \int_0^1 \sum_{k=1}^{(M+N)x} y_k dx \tag{4}$$

So are the two departments.

$$G_1 = 1 - \frac{2}{Y_1} \int_0^1 \sum_{k=1}^{xM} y_k dx \tag{5}$$

$$G_2 = 1 - \frac{2}{Y_2} \int_0^1 \sum_{k=M+1}^{M+xN} y_k dx \tag{6}$$

The income ratio function definition of the two sectors has been taken into account

Now, the Gini coefficient relationships between the whole society and the two departments are considered in accordance with the above definitions and Figure 1. In fact, The centre result of this paper follows from the straightforward calculations as

$$\begin{aligned} G &= 1 - \frac{2}{Y} \int_0^1 \sum_{k=1}^{(M+N)x} y_k dx \\ &= 1 - \frac{Y_1}{Y} (1 - G_1) - \frac{Y_2}{Y} (1 - G_2) + \frac{2}{Y} \int_0^1 F(x) dx \end{aligned} \tag{7}$$

The individual total income proportions  $Y_1/Y$  and  $Y_2/Y$  of two departments to the entire society income are the respective weighted Gini coefficient contributions to the whole society Gini coefficient. The left third plot function in formula (7)

$$F(x) = \sum_{k=(M+N)x+1}^{M+xN} y_k - \sum_{k=xM+1}^M y_k = \sum_{k=xM+1}^M (y_{k+xN} - y_k) \geq 0 \quad (8)$$

Represents the calculated income difference within the width of  $(1-x)M$  in the two different departments, in which the first term is the sum of the income in the interval  $[(M+N)x+1, M+Nx]$  in the first sector; the second term is the sum of the income in the interval  $[Mx+1, M]$  in the second sector. It is positive because the income of the second sector is more than that of the first sector ( $y_{k+xN} \geq y_k$ ),  $F(x) \geq 0$ .

The expression (8) above is one of the core conclusions in this paper, which tells Gini coefficient of the whole society is the weighted average of both the urban and rural sectors' Gini coefficients to their income respectively (depending on the respective proportion of the urban and rural sectors' total income to the whole society income) with a volume greater than zero. Its economic meaning is very clear: the income distribution fairness of the whole of society depends not only on the respective income distribution fairness within the urban and rural departments, but also on their respective income ratio or the income distribution fairness level between the two departments if a community is composited by two large-scale sectors. In other words, the whole community will be unfair if the allocation between the two departments is not fair no matter how fair social distribution within the two departments is, which has the same meaning that the inner fairness of the two departments can not determine the fairness of the whole community. Particularly, the whole community's Gini coefficient can be significant (such extreme will be analyzed in detail in

the next section) due to the volume greater than zero even though the internal income distribution is of average within the urban and rural sectors ( $G_1 = G_2 = 0$ ). Statistically, the phenomenon is the so-called "Simpson's Paradox"<sup>1</sup>. The main reason leading to Simpson's Paradox is if the difference between the groups is large enough, which is profoundly indicated from distribution relationships between the rural and urban departments in this paper. Moreover, we can see that the Gini coefficient of the whole society is correlated with the income gap between urban and rural areas positively, that is, narrowing the urban-rural income gap is the necessary condition to narrow society's Gini coefficient, and the Gini coefficient problem of the whole society can not be solved only considering the income distribution fairness within the urban and rural departments respectively.

From the discussion above, we can see that if there is a sector (such as China's rural sector) with the majority population of the total population in a economical society,, even if the internal unfairness is in a low level, its impact to the unfairness of the whole society is quite big due to the large revenue weight. Therefore, the crux of the problem lies not in whether it is relatively fair in the rural sector with the majority of the total population but in whether the income distribution between the rural sector and the urban sector (a minority of the total population) is fair. In fact, because the society is an organic whole, its fairness is

---

1 The so-called Simpson's Paradox is such a phenomenon wherein the successes of groups seem reversed when the groups are combined. This phenomenon was widely concerned after the publication of EH Simpson's papers in 1951. There is an example usually cited to illustrate this phenomenon. We discuss if there exist gender discrimination in the new semester enrollment of Law and Business School. Within the law school, a higher proportion of females are admitted (8/53=15.1% of the males were admitted compared to 51/152=33.1% of the females). A higher proportion of the females are admitted to the business school as well (201/251=80.1% of the males were admitted compared to 92/101=91.1% of the females). But when look at two schools wholly, we observe quite a different picture. The data indicate that, overall, a lower proportion of women are admitted into the university (68.8% of the males compared to less than 56.5% of females).

not only determined by the respective common fairness of the two different departments, but by the fairness between the two departments. Because the population of the two departments is large enough, their sense of injustice is significant, especially the low-income farmers are nearly 60 percent of the total population. Of course, this sense of fairness is not just a perception, but has its real objective bases. It will go as the assertion: farmers' sense of unfairness is not so prominent due to the urban and rural dual structure. According to this assertion, it seems that the urban and rural dual structure will shield most of the information about the gap between urban and rural areas, thus, farmers are still quite satisfied to live on. However, the fact is not the case. Urban and rural dual structure has not only exacerbated inequality in urban and rural areas (enabling farmers and urban residents do not receive the same national treatment), but also deepened peasants' sense of unfairness. Because though the dual structure can prevent farmers' identity changes (changes to citizens), the dissemination of information is not prevented. Apart from the fact of the rapid transmission of television and network, the huge gap between urban and rural areas is felt thoroughly and household disseminated only by the huge peasant worker tide annually surging between the urban and rural areas. And, what's more, the gap will be enlarged consciously or unconsciously because of the imagination. According to a comparative analysis on the living conditions of the urban and rural residents in Zhejiang Province, the happy situations of urban residents are better than the rural residents, and the difference between the two achieves a significant level. As the richest farmers' province for 19 years, the development gap between urban and rural areas is so obvious in Zhejiang, and the gap between urban and rural areas in the other provinces can be imagined naturally.

In order to deepen the understanding of the core conclusion in this paper, a special but closely contacted with the reality situation will be analyzed below:  $G_1 = G_2 = 0$ . It is easy to obtain the Gini coefficient of the whole society

$$G = \frac{Y_2}{Y} - \frac{N}{M+N} = \frac{M}{M+N} - \frac{Y_1}{Y} \quad (9)$$

Performs as the differences between the population ratio and the income ratio of a department (G has two expressions as showed in the above equation).

The analysis of the above example shows once again that the fairness of the vast majority is the fairness of the whole society. It is not consistent with the basic principles of economics theories if a society is divided to several parts and the income distribution fairness is calculated as the so-called "one country, two measure" considering the separate fairness within each department. The Gini coefficient in economics is a whole notion measured by the income, that is, if we want to calculate the Gini coefficient of a country (or a province), it is necessary to arrange the residents in the country (province) together only

on the sort of income, rather than calculate it respectively after first divisions on other criteria such as the urban and rural areas, regions or industries. If a certain point is on the case, for an extreme example: the division will come into a family at last. In this way, the Gini coefficient will be the weighted average after calculation according to families and the results will undoubtedly show complete fairness of the whole community regardless the fact how unfair the society is. It will be apparently senseless if the Gini coefficient is obtained in this way.

In addition, it is contrary to the facts that the China's Gini coefficient is treated as the so-called "national conditions discount" under the pretext of the dual structure of China. In fact, the phenomenon of China's urban and rural dual economy itself reveals the serious inequalities in the economy, income, interest fields between urban and rural areas. And it is this serious inequality which increases its Gini coefficient which highlights the alarming of the Gini Coefficient. If the "urban and rural dual economy" widening the income gaps is regarded as a reason for "the tolerance of the widening income gap", or the Gini coefficient is made incredible and needs make a discount because of the existence of the urban and rural structure. Accordingly, to find out the causes of the increased Gini coefficient and each of them is to be discounted. It is clear that the Gini coefficient after the ultimate discount would be unable to measure the extent of inequality, and thus it will be not the Gini coefficient any more.

In summary, the Gini coefficient is an effective indicator measuring the income distribution gap of a whole society. There should be no "discount" on the international used Gini coefficient just because of the existence of China's urban and rural economy dual structure, that is, the Gini coefficient of China's rural and urban sectors is calculated dividedly and the Gini coefficient of the entire society is the simply sum of the results. It is not consistent with the calculation rules of the Gini coefficient, and so that the data is impossible to reveal the income distribution fairness of the whole society. As urban and rural areas are two departments with interactions in a social, if we want to get the Gini coefficient of the whole society, it should be described by the extended Lorenz curve on the basis of the Gini coefficient principles rather than the simple sum of the respective Gini coefficient of the urban and rural areas.

In this paper, the Lorenz curve is extend, and the Gini coefficient of the two interacting urban and rural departments and the relationships between the Gini coefficient of the whole society and that of the two departments are described in detail using the generalized Lorenz curve. The conclusion is that there exists social injustice based on two entirely fair departments. This conclusion is fully consistent with the "Simpson's Paradox" principle. Based on this conclusion, it is further pointed out that the income distribution inequality can not be overlooked even if the internal income distributions within the urban and rural departments are relatively equal. Therefore, effective measures, such as increasing financial support in the rural areas, accelerating the transfer of labor force and population

flows, improving the rural financial system, should be taken actively by the government to narrow the income gap between urban and rural areas so as to improve the income distribution inequalities between urban and rural areas.