

EDUCATIONAL INEQUALITY AND EDUCATIONAL POVERTY. THE CHINESE CASE IN THE PERIOD 1975-2004

Donatella Saccone

Department of Economics "S. Cogneetti de Martiis"
University of Turin

Abstract

Together with the economic transition, in China the return to education and the skill premium increased; this phenomenon was deeply demonstrated by a large number of studies identifying it with one of the main cause of the increasing inequalities at national, regional and sector level. For this reason, these studies underlined the relevant role of education in influencing the future evolution of the income distribution.

Our study aims to analyze the evolution of educational inequality and educational poverty over the period 1975-2004 by using traditional inequality indicators, but adapting them to some features of education distribution. In particular, we exploit the Theil index decomposition properties to distinguish between distributive improvements due to changes in basic education diffusion and improvements related to all the educational levels. The analysis is carried out by using the China Health and Nutrition Survey provided by the Carolina Population Center for the period 1989-2004. In spite of its incomplete geographical coverage (only 7 Chinese provinces were considered), this survey allows us to follow the evolution of educational inequality and poverty over 15 years. Moreover, we also use other data sources representative at national level: the Barro-Lee dataset, providing observations from 1975, and the data published in the China Statistical Yearbooks.

The results show that educational inequality indicators constantly decreased not only at national level but also within each province and area. However, the Theil index decomposition shows that the decrease of educational poverty and the reduction of its depth played a relevant role in improving the whole education distribution. On the contrary, with regard to people accessing to education, educational inequality indicators show a tendency to increase over the considered period; this reveals that the strongest changes and distributional progress concerned the lowest part of the education distribution, while inequalities among educated people didn't improve.

Key words: Human Capital; Educational Inequality; Educational Poverty; China.

JEL classification: C43; D63; I21; I32; O15.

1. Introduction

When we are dealing with the concept of inequality, we usually refer to the income or expenditure inequality; the same is true for the notion of poverty. However the concepts of inequality and poverty can be applied and adapted to study the distribution of other assets and, in particular, of education. Although there exist some relevant works on this issue¹, the analysis of education distribution is still scarce if compared with the studies on income distribution. Nonetheless, it is highly important to understand the dynamics related to the knowledge diffusion and distribution, above all in the contemporary society. In fact, in the knowledge and information era education seems to be one of the most significant assets largely influencing how an individual can interact and take part in the economic, social and political life².

However, income distribution and education distribution differ in some important regards. Educational inequality changes are characterized by trends differing from the income distribution dynamics in terms of intensity and stability. In fact, the dynamics of education distribution are slower but also more stable and continuous; this is because the changes in human capital can involve one or more generations. Moreover, once the education distribution becomes more even, it seems to be a permanent change in most cases; Thomas et al. (2002) calculated educational inequality indicators for 140 countries in the period 1960-2000, finding a constant increase of equity for the most part of analyzed cases. Finally, distributive policy for income and education -even if partially linked- may differ. In particular, the education distribution can be improved by a specific policy, i.e. by introducing compulsory education or increasing its length. This policy is often implemented in DCs to reduce the illiteracy rate and the schooling poverty; it acts above all on the lowest part of education distribution. However, in this case the decreasing of educational inequalities is due to a larger access to basic education but it is less related to the access to secondary or post-secondary education.

In the light of these differences, our analysis aims to use traditional inequality indicators, but adapting them to some features of education distribution. In particular, we will exploit the Theil index decomposition properties to distinguish between distributive improvements due to changes in basic education diffusion (for example due to compulsory education laws, measures decreasing illiteracy and educational poverty etc.) and improvements related to all the educational levels. Some studies decomposed the Theil index by using micro-data; however, a specific and explicit formula to decompose educational inequality was never proposed; this formula could be useful when we have no individual and continuous observations but only the population shares for each educational level. At first, we will propose the general formula to decompose the Theil index of education distribution. Then we will introduce the concept of educational poverty and we will see how it can be related to educational inequalities by exploiting the Theil index decomposition properties; in the context of DCs this methodology is particularly useful to understand the dynamics of educational inequality among the poor in education, among educated people and between these two groups.

¹ Ram (1990), Lopez et al. (1998), Checchi (2001), Thomas et al. (2001), Checchi and García-Peñalosa (2004), Lim and Tang (2006), Lin (2007) are some of the recent contributions to this subject.

² On the importance of education in the contemporary society, see Perez (1994 and 2001), Hodgson (2001) and Stewart (1996).

Finally, we will apply the methodology to the Chinese case in the period 1975-2004 by studying the evolution of educational inequalities and the dynamics of schooling poverty in relation to the income distribution.

2. Methodology

2.1. Indicators of education inequality and Theil index decomposition

Unlike the analysis of income inequality, the study of educational inequality is not much widespread and, in most cases, it was carried out cross-country by using the Barro-Lee dataset³. In these studies various indicators and methods measuring educational inequalities were proposed; here we will adopt the education Gini coefficient proposed by Checchi (2001) and the education Theil index suggested by Thomas, Wang and Fan (2001). The education Gini coefficient ($Gini_{ed}$) is defined in the following way:

$$Gini_{ed} = HC_n + \frac{1}{HC} \sum_{i=2}^N \sum_{j=2}^N p_i p_j |n_i - n_j|, \text{ where}$$

N = number of education levels

p_i (p_j) = share of population with the education level i (or j)

n_i (n_j) = length of the education level i (or j)

\overline{HC} = average years of education

HC_n = share of uneducated population.

Moreover, the education Theil index is given by:

$$Theil_{ed} = \sum_{i=1}^N p_i \left(\frac{n_i}{\overline{HC}} \right) \ln \left(\frac{n_i}{\overline{HC}} \right).$$

On the contrary, a specific formula to decompose the education Theil index by groups was never expressed. Some studies indeed decomposed it without formally making explicit a formula for the case of education distribution. However, this formula could be useful when we have the population shares for each education level. We propose the following methodology to decompose the index by groups:

$$Theil_{ed} = \sum_{i=1}^J \left(\frac{A_j}{A} \right) T_j + \sum_{i=1}^J \left(\frac{A_j}{A} \right) \ln \frac{A_j / A}{P_j / P}.$$

³ Ram (1990), Lopez et al. (1998), Checchi (2001), Thomas et al. (2001), Checchi and García-Peñalosa (2004), Lim and Tang (2006) analyzed the education distribution by using this dataset.

J is the number of groups by which the index is decomposed. A can be interpreted as the total sum of years spent studying by the population; in other words, $A = \sum_{i=1}^N P_i n_i$, where P_i is the number of individuals with the education level i and n_i is the length (expressed in years) of that level. The ratio A_j / A represents the total number of years spent studying by the group j (A_j) as a share of the total number of years spent studying by the whole population (A); this ratio can be interpreted as the *education share* of the group j . Analogously, P_j / P represents the *population share* of the group j . Finally, T_j is the education Theil index referred to the group j and calculated by the formula previously showed. In other words, the first addend represents the *within-component* of the Theil index, i.e. the part of inequality generated by within-group disparities; on the contrary, the second addend expresses the *between-component* depending on between-group inequalities.

Since $\frac{A_j}{A} = \frac{P_j \overline{HC}_j}{PHC}$, the formula can also be expressed in terms of average years of education:

$$Theil_{ed} = \sum_{i=1}^J f_j \frac{\overline{HC}_j}{\overline{HC}} T_j + \sum_{i=1}^J f_j \frac{\overline{HC}_j}{\overline{HC}} \ln \frac{\overline{HC}_j}{\overline{HC}},$$

where f_j is the *population share* of the group j .

2.2. Educational inequality and educational poverty: a possible decomposition

In the field of education, another concept scarcely analyzed is that of educational poverty. However, the study of educational poverty is important above all when DCs are the object of analysis; in fact, this phenomenon can concern a significant population share in these countries.

When we are dealing with educational poverty, we have to identify a threshold below which an individual is poor in education. This point was examined by Checchi (1996), who showed that using absolute rather than relative thresholds is more appropriate for the analysis of educational poverty. In other words, Checchi proposed to use the years of compulsory education as a threshold, on condition that they have been completed by the 60% of compelled people. However, in the case of DCs, there could be cases and historical periods in which there is no compulsory education or this last has not been completed by the 60% of compelled population. In these cases, we can appeal to other criteria of specification which provide interesting information even though more arbitrary. Some examples of threshold are: the compulsory education (without condition of accomplishment), the primary education, the time necessary to become literate (in general three years or more), and so on.

When a threshold is identified, we propose to use the traditional indicators of poverty usually applied to the income distribution analysis. In particular, the first

indicator is the Poverty Headcount Ratio, H , defined by the poor population (P_p) as a share of total population (P):

$$H = \frac{P_p}{P} .$$

Moreover, in the case of DCs it is particularly important to measure the educational poverty depth; this information can be obtained by calculating the Average Poverty Gap:

$$APG = \left(\frac{1}{P} \right) \sum_{i=1}^p \frac{(z - n_i)}{z}$$

where z expresses the educational poverty threshold and n_i represents the number of years spent studying by the poor individual i .

Similar to the case of income distribution, also in the case of education distribution there is a relationship between inequalities and poverty; however, the two cases differ in some regards. In the case of income distribution, a decrease of poverty doesn't necessarily cause a decrease of inequalities. In the case of education distribution the reasoning is different; in fact, while the income of upper deciles can increase without limits, there usually is an upper bound to the education level that each individual achieves. For this reason, the range of education growth is larger for the lowest part of the distribution; then, in the case of education distribution, a decrease of poverty is usually reflected in a decline of inequality indicators. However, this decline can be generated just by what happens below and around the threshold, but it is not sure that it is also due to a more even access to all the education levels- including the highest levels. In fact, this information can be misleading if not combined with a deeper analysis.

In other words, we are stating that the study of inequality is not sufficient in the context of education; on the contrary, it has to be combined with the study of educational poverty and a series of methodologies making possible to distinguish between distributive improvements due to a decrease of educational poverty and improvements related to a larger diffusion of all the educational levels. The following methodology arises from this statement and is based on the Theil index decomposition properties.

Let's divide the population, P , in two groups: 'the poor' and 'the not poor' in education. The first group is then represented by the population share below the educational poverty threshold, P_1/P , while the second one by the population share above this threshold, P_2/P . The education Theil index is given by:

$$Theil_{ed} = \frac{A_1}{A} T_1 + \frac{A_2}{A} T_2 + \frac{A_1}{A} \ln \frac{A_1/A}{P_1/P} + \frac{A_2}{A} \ln \frac{A_2/A}{P_2/P} .$$

Since P_1/P merely is the Poverty Headcount Ratio ($P_1/P=H$), we can rewrite the education Theil index in the following way:

$$Theil_{ed} = \frac{A_1}{A} T_1 + \frac{A_2}{A} T_2 + \frac{A_1}{A} \ln \frac{A_1/A}{H} + \frac{A_2}{A} \ln \frac{A_2/A}{(1-H)}$$

or in terms of average years of education:

$$Theil_{ed} = H \frac{\overline{HC_1}}{\overline{HC}} T_1 + (1-H) \frac{\overline{HC_2}}{\overline{HC}} T_2 + H \frac{\overline{HC_1}}{\overline{HC}} \ln \frac{\overline{HC_1}}{\overline{HC}} + (1-H) \frac{\overline{HC_2}}{\overline{HC}} \ln \frac{\overline{HC_2}}{\overline{HC}} .$$

As we have already seen, the first two addends represent the within-component, while the other two addends the between-component. In particular, the second addend expresses how much part of total inequality is explained by what happens above the educational poverty threshold; by studying the temporal change of this component, we can see how the inequality generated by differences in the access to the highest education levels evolves.

Since we are also interested in the dynamics of educational inequality, we can calculate how much each component contributes to the Theil index change between two periods:

$$\Delta Theil_{ed} = \Delta \left(H \frac{\overline{HC_1}}{\overline{HC}} T_1 \right) + \Delta \left((1-H) \frac{\overline{HC_2}}{\overline{HC}} T_2 \right) + \Delta \left(H \frac{\overline{HC_1}}{\overline{HC}} \ln \frac{\overline{HC_1}}{\overline{HC}} \right) + \Delta \left((1-H) \frac{\overline{HC_2}}{\overline{HC}} \ln \frac{\overline{HC_2}}{\overline{HC}} \right).$$

In this way, we can obtain some relevant information making more complete the temporal analysis of education distribution:

- the first addend suggests how much part of the Theil index change is due to dynamics within the group of ‘the poor’ in education; these dynamics, in turn, depend on the changes in educational poverty diffusion, on inequalities among ‘the poor’ in education, and on the relative distance of the average years of education of this group with respect to the average years of the whole population;
- analogously, the second addend expresses how much part of the Theil index change is related to dynamics within the group of ‘the not poor’ in education; these dynamics, in turn, originate from the changes in the ‘not poor’ population, from inequalities within this group, and from the relative distance of the average years of education of this group with respect to the average years of the whole population;
- by summing the first and the second addends, we obtain the component of the Theil index change generated by within-dynamics;
- the third (the fourth) addend represents the part of the index variation due to changes in the relative distance between the average years of the first (second) group and the average years of the whole population;
- the sum of the last two addends expresses how much part of inequality change is due to dynamics between the two groups, i.e. to changes of the between-component.

In the following part, we will show an application of this methodology to the Chinese case.

3. An application to the Chinese case in the period 1975-2004

Studying the evolution of Chinese education distribution is interesting for a number of reasons. First of all, since the end of the '70s China showed deep and sharp changes also regarding the education levels of people: starting from a very high illiteracy rate, after 20 years of reforms basic education largely spread. Governmental policy led and stimulated this change: in the 1986 the *Compulsory Education Law for the People's Republic of China* was introduced; this law provided for 9 years of compulsory education and was coupled with a series of efforts to disseminate basic education also in the rural and poorest areas. Moreover, many studies analyzing the Chinese income distribution underlined the importance of education in determining inequalities among individuals, areas and provinces; in other words, education was seen as the key to building the future income distribution. In this regard, Fang et al. (2002) stated "*investment in human capital is key to long-term improvements in welfare for all*". Also Heckman (2005) declared "*human capital is the asset that ultimately determines the wealth of China. Fostering access to education will reduce inequality in the long run*", while Benjamin et al. (2005) underlined "[...] *the important role played by education in both urban and rural areas. Probably no other single factor will be most closely tied to how the fruits of future growth are shared*".

In the following part we will try to understand the dynamics of the Chinese education distribution and, in particular, whether the improvements concerned all the educational levels or were limited to basic education. Moreover, we will compare the evolution of income and education distribution to verify if a common trend is identifiable. First of all we will present the data and their characteristics and, then, we will show the results of our analysis; after describing the general trend of educational inequality over the considered period, we will analyze its relationship with the educational poverty by using the methodology above illustrated.

3.1. Data description

The main source on which we will base our analysis is the *China Health and Nutrition Survey* (CHNS). This is a multistage random cluster sample survey carried by a group of social scientists and biomedical researchers under the control of the Carolina Population Center, the National Institute of Nutrition and Food Safety, and the Chinese Center for Disease Control and Prevention. The survey was collected in various years (1989-1991-1993-1997-2000-2004) and the Carolina Population Center provides a longitudinal database making easier the inter-temporal and inter-spatial comparability. Even though its focus is on health issues, the survey also includes the number of education years completed by each interviewed individual and a series of variables regarding several individual characteristics (age, province of residence, area of residence, etc.). However, the survey covers just 9 Chinese provinces: Liaoning, Heilongjiang, Jiangsu, Shandong, Henan, Hubei, Hunan, Guangxi, Guizhou. Moreover, since the provinces of Liaoning and Heilongjiang are not represented in every year, we will limit our analysis to the remaining 7 provinces, which are marked in the map

(figure 1)⁴. Anyway, the covered provinces differ with respect to geography, economic characteristics, social indicators, and public resources. Since the richest and poorest provinces were excluded, the survey could underestimate the level of heterogeneity characterizing the country (Benjamin et al., 2000).

Figure 1. Map of China.



Despite of these limits –first of all the incomplete geographical coverage- the CHNS possesses some useful characteristics for our purpose. On the one hand, it allows us to analyze the education distribution by using microdata; this is particularly important for the Chinese case, in which the lack of household surveys highly limits the carrying out of deep analysis on income and education distribution. On the other hand, given its temporal coverage, the survey makes us able to jointly study the evolution of

⁴ The survey is conducted over a 3-day period on a sample of about 4400 households (19000 individuals). However, the longitudinal database is smaller: excluding the provinces of Liaoning and Heilongjiang, we have observations for about 3300 households (9300 individuals). For details see the website of the Carolina Population Center: www.cpc.unc.edu/projects/china/data.

education distribution and of income distribution over a period of 15 years. Moreover, each covered province has a population which is larger than the population of many DCs; for this reason, although the conclusions cannot be generalizable to the whole country, they are generalizable to a significant quantity of individuals (Benjamin et al., 2000).

The CHNS provides the number of completed years of education for each individual. In this way, we can deal with continuous rather than discrete observations and, then, obtain more reliable estimates of inequality indicators. Moreover, also information on gender, province and area of residence (urban or rural) is available for each individual; this allows us to carry out a decomposition of inequality by gender, provinces and areas. We will focus just on individuals aged 15 and above, who were not enrolled when the survey was conducted.

In addition, this study will be combined with other two data sources; this is useful in order to compare the results emerging from these different sources and test their reliability. The first source is the Barro-Lee dataset (Barro and Lee, 2000)⁵; it collects data on educational attainment by levels for two age groups (over-15 and over-25) and for 138 countries over the period 1960-2000 (observations are at 5-year intervals). The considered levels of education are: no schooling, primary education, secondary education and post-secondary education (the last three levels are divided in two categories of accomplishment: complete and incomplete)⁶. We will use the data concerning the over-15 population; for the Chinese case, they are available for the period 1975-2000. By using this source, we are able to estimate the indicators of educational inequality at national level.

Moreover, the China Statistical Yearbook (CSY) reports educational attainments of the over-6 population by gender and for five education levels: no schooling, primary education, secondary education-1° cycle, secondary education-2° cycle, post-secondary education. The CSY data are based on the Population Sample Survey, conducted on a sample of about 1 per cent of total population. Given their national coverage, the data are also divided by provinces. This allows us to estimate the indicators of educational inequality at national level, and by provinces and gender; moreover, we are able to carry out a decomposition of inequality by groups (provinces and gender). We will use the data for three different years: 1997-2000-2004.

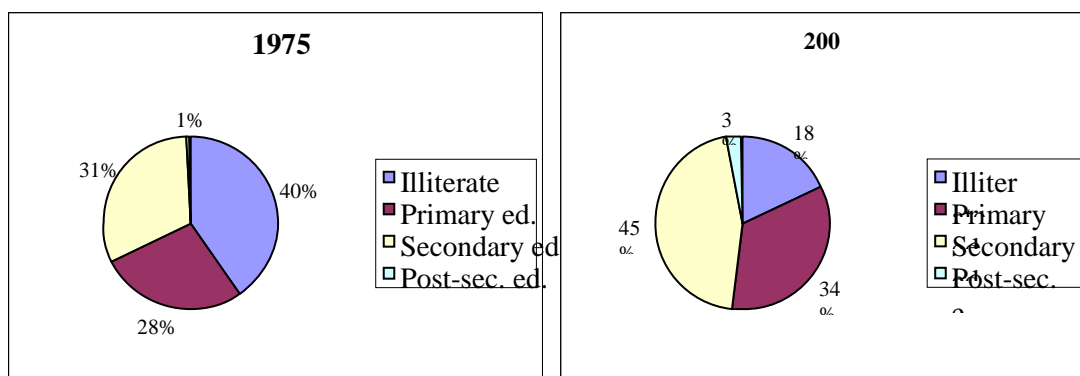
3.2.1. *The Chinese education distribution in the period 1975-2004*

Over 25 years the quantity and quality of China's human capital stock dramatically changed (graph 1). In 1975, the education distribution was highly polarized: 40% of population was illiterate, while one third of people had secondary education (incomplete or complete). Over the period 1975-2000 the percentage of illiterate people significantly declined from 40% to less than 20%; at the same time, the percentage of people with primary and, above all, secondary education increased. This was generated by the introduction of compulsory education providing for 9 years of schooling: 6 years of primary education and three years of secondary education (1° cycle). However, in 2000 the share of population with high education was still low.

⁵ The data are available on line: www.cid.harvard.edu/ciddata/ciddata.html.

⁶ We will assume that the length of each *incomplete* level of education is equal to half the length of the *complete* level.

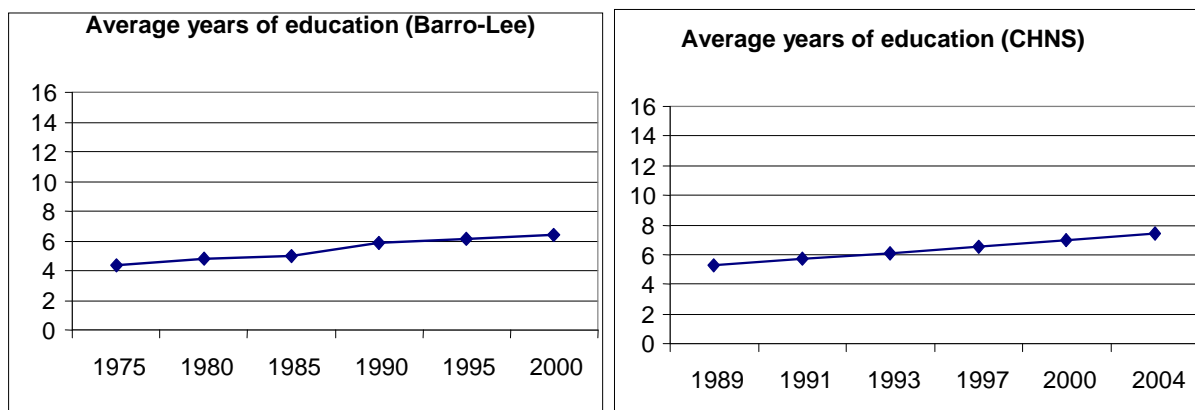
Graph 1. Population shares by education levels (1975 and 2000)



Source: Barro-Lee dataset (2000).

The educational improvement can also be typified by looking at the trend of the average years of education. Despite of some differences in absolute values, both the Barro-Lee dataset and the CHNS data show a constant growth in the average years of education (graph 2). The former source reveals that over the period 1975-2000 the average years of education increased from 4.4 (incomplete primary education) to 6.4 (beyond complete primary education).

Graph 2. Trend in the average years of education

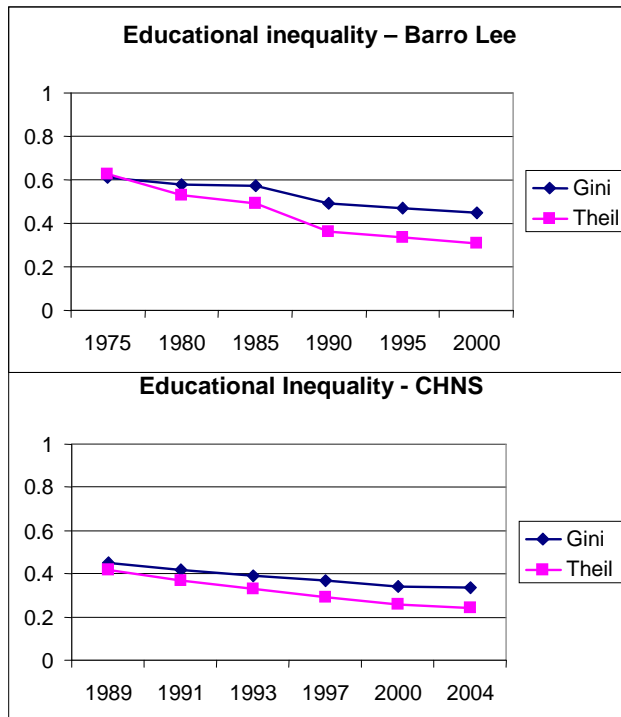


Source: Barro-Lee dataset and our calculations on the CHNS data

These changes were reflected in an unequivocal drop of educational inequality indicators; this point is confirmed by both data sources, although there exist some estimation discrepancies due to the different geographical coverage (graph 3). In this case the estimates based on the Barro-Lee dataset are more reliable, because referred to the whole country; from them it emerges that the education Gini coefficient declined from a very high value of 0.61 to 0.45 over the period 1975-2000, showing a reduction of 27%.

It is worth comparing the trend of the education Gini coefficient to that of the education Theil index. At the beginning of the period, the value of this latter was similar to the Gini coefficient value; however, over the analyzed period the distance between these two values constantly widened. This is because the Theil index is much more sensitive to values near to 0: the increasing distance between the two indicators then reflects the dramatic fall in the share of illiterate population.

Graph 3. Evolution of the educational inequality indicators.



Source: our calculations based on the Barro-Lee dataset (2000) and the CHNS data (various years).

Educational inequality indicators decreased not only at national level. They also declined in each of the 7 CHNS provinces; within both urban and rural sectors without remarkable differences in trends and absolute values; within each gender group (see table 1, graph 1 and graph 2-3 in appendix respectively). In this last case, however, we have to notice some important differences. At the beginning of the period, educational inequalities among males were already low, while they were very high among females (the Gini coefficients were around 0.4 and 0.6 respectively). Although inequalities declined within both groups over the period 1989-2004, in 2004 they still remained at a very high level for females; at the end of the analyzed period, the Gini coefficient value for females was higher than that for males at the beginning of the same period.

With respect to the within-group educational inequality, the trend of the between-group inequality was less stable and generalizable. In the table 1, we reported the values of a sort of educational gap between sectors and between genders; for each of them, this gap was expressed by dividing the average years of education of one group by those of the other group (urban sector/rural sector and males/females).

Table 1. Educational inequalities between sectors and genders.

	Average years of education urban sector/rural sector	Average years of education male/female
1989	1.2169	1.4868
1991	1.2119	1.4495
1993	1.1964	1.4438
1997	1.2022	1.3875
2000	1.1802	1.3709
2004	1.2384	1.3495

Source: our calculations based on the CHNS data.

Although the between-sector gap does not show any clear temporal trend, it reveals how much the educational difference between urban and rural areas is large: in 2004 an individual residing in the former studied on average 24% more than an individual residing in the latter. On the contrary, the evolution of the between-gender inequality is clearer since the educational gap between males and females constantly decreased: while in 1989 a male studied about 50% more than a female, in 2004 this percentage fell to 35%. Finally, the table 2 in appendix shows the spatial educational inequality, measured by the standard deviation of the provincial average years of education. In this case, the analyzed period is shorter (1997 and 2004) since for this purpose we had to use the CSY data. Although the standard deviation slightly decreased over this period, educational spatial inequalities are still deep. In 2004, the Northern and North-eastern provinces showed an average level of education similar to that of many advanced countries (the province of Beijing was characterized by an average level of education near to the South Korean one); on the other hand, the average years of education were lower in the Eastern and central provinces, although in this regard the poorest provinces were the Western ones.

When we decomposed the educational inequality by groups – provinces, areas and gender - we obtained an interesting picture (table 3 and 4 in appendix). The within components have a predominant importance in determining the whole educational inequality, beyond 95%. This point fits into the results of existing literature which analyzed the Chinese income distribution; in fact, some recent studies reassessed the importance of the between dynamics and stressed the major role played by the within-group income inequality (Lee, 2000 ; Benjamin et al., 2000 and 2005; Sicular et al., 2007). Our study confirms these findings, by showing that also from an educational point of view the most part of inequality is generated by disparities existing within each province, each area and each gender.

3.2.2. Educational inequality and educational poverty

Although the Chinese progress in education was evident and unambiguous, we have to not confuse the decrease of the illiteracy rate with an improvement generalized to all the education levels. For this reason, it is useful to apply the previously presented methodology. First of all, we have to define an educational poverty threshold for the Chinese case. As we have already illustrated, in China the compulsory education was introduced just in 1986 and required 9 years of schooling to start at the age of 7; calculating the percentage of compelled people who accomplished it is possible just

from the year 1995. Since the CHNS was collected in 1989-1991-1993-1997-2000-2004, we have just 3 years after this date and a very low percentage of compelled people (people born from 1979 onwards) in the whole sample population. As a consequence, the limited number of observations doesn't allow us to get efficient estimates of the percentage of compelled people who completed compulsory education. In the light of this, we chose an educational poverty threshold equal for all individuals, constant over time and with an institutional value: this is the compulsory education, i.e. 9 years of schooling.

Table 2: Decomposition of the education Theil index by "the poor" and "the not poor" in education – people aged 15-40.

Threshold=9	1989	1991	1993	1997	2000	2004
1. THEIL	0.18966	0.15131	0.12365	0.08401	0.06523	0.06939
2. GINI	0.29252	0.25765	0.23008	0.19339	0.17235	0.18301
3. \overline{HC}	6.9605	7.5018	7.8793	8.5017	8.9965	9.4241
4. $P_1/P = H$	0.5829	0.5006	0.4453	0.3612	0.2800	0.2346
5. $P_2/P = (1-H)$	0.4171	0.4994	0.5547	0.6388	0.7200	0.7654
6. APG	0.28288	0.23300	0.19691	0.14448	0.10777	0.09754
7. \overline{HC}_1	4.6326	4.8111	5.0203	5.4000	5.5357	5.2578
8. \overline{HC}_2	10.2143	10.1990	10.1745	10.2555	10.3422	10.7010
9. $\overline{HC}_1 / \overline{HC}$	0.6656	0.6413	0.6371	0.6352	0.6153	0.5579
10. $\overline{HC}_2 / \overline{HC}$	1.4675	1.3595	1.2913	1.2063	1.1496	1.1355
11. GINI 1	0.34165	0.31572	0.29042	0.23605	0.21133	0.21061
12. THEIL 1	0.27167	0.23883	0.20889	0.14159	0.11444	0.11736
13. GINI 2	0.07948	0.08069	0.08007	0.08542	0.09138	0.10950
14. THEIL 2	0.01222	0.01284	0.01279	0.01445	0.01666	0.02298
15. THEIL W	0.11289	0.08539	0.06843	0.04361	0.03350	0.03533
16. THEIL B	0.07677	0.06592	0.05522	0.04039	0.03172	0.03406
17. % W	59.52	56.43	55.34	51.91	51.36	50.92
18. % B	40.48	43.57	44.66	48.08	48.63	49.08

Source: our calculations based on the CHNS data.

The table 2 reports the main results of this analysis. The first three rows show the education Theil indexes, the Gini coefficients and the average years of education for people aged 15-40; since old generations are just partially involved in new educational opportunities, we chose to consider only young people. As we see in the case of the total sample, both indexes decrease in the period 1989-2004. In the two following rows

we can see the population shares of the two groups, “the poor” and “the not poor” in education respectively; H obviously expresses the Education Poverty Headcount Ratio. These results reveal that the share of “the poor” in education constantly decreased over the analyzed period; this share moved from 0.58 to 0.23. Moreover, the values of the Average Education Poverty Gap (row 6) confirm that not only educational poverty but also its depth diminished; in other words, the average distance between the schooling years of “the poor” and the educational poverty threshold narrowed. The rows 7 and 8 report the temporal trend of the average years of education for the two groups: average education increased for both groups, but this increase was relatively higher for “the poor” (13.5% with respect to 4.8% for “the not poor”).

It is interesting to look at the trend of educational inequality indicators within each group (rows 11-14) because the results confirm our doubts. Within the group of “the poor” in education, indicators show a constant decrease of educational inequality although they remain at a significant level because of a still important presence of illiterate people. On the contrary, with regard to “the not poor” in education, inequality indicators –even though at a lower level- showed a tendency to increase. This reveals that the strongest changes and distributional progress concerned the lowest part of the education distribution, while inequalities among people accessing to education didn’t improve.

Finally, the last four rows report the results of the decomposition. Looking at the percentages, it is clear that the polarization between “the poor” and “the not poor” in education intensified; while in the 1989 the most part of educational inequality among young people was explained by the within component, in the 2004 each of the two components was responsible for one half of total inequality .

We also decomposed the Theil index change; the table 3 presents the main steps and the results.

Table 3: Decomposition of the educational Theil index change by " the poor" and " the not poor" in education.

	$H \frac{\overline{HC_1}}{HC} T_1$	$(1-H) \frac{\overline{HC_2}}{HC} T_2$	$H \frac{\overline{HC_1}}{HC} \ln \frac{\overline{HC_1}}{HC}$	$(1-H) \frac{\overline{HC_2}}{HC} \ln \frac{\overline{HC_2}}{HC}$	THEIL
1989	0.1054	0.0075	-0.1579	0.2348	0.1897
2004	0.0154	0.0200	-0.0764	0.1104	0.0694
Δ 1989-2004	0.0900	-0.0125	-0.0815	0.1243	-
Contribution (%)	74.8	-10.4	- 67.8	103.3	100
	% WITHIN		% BETWEEN		
	64.4		35.6		100

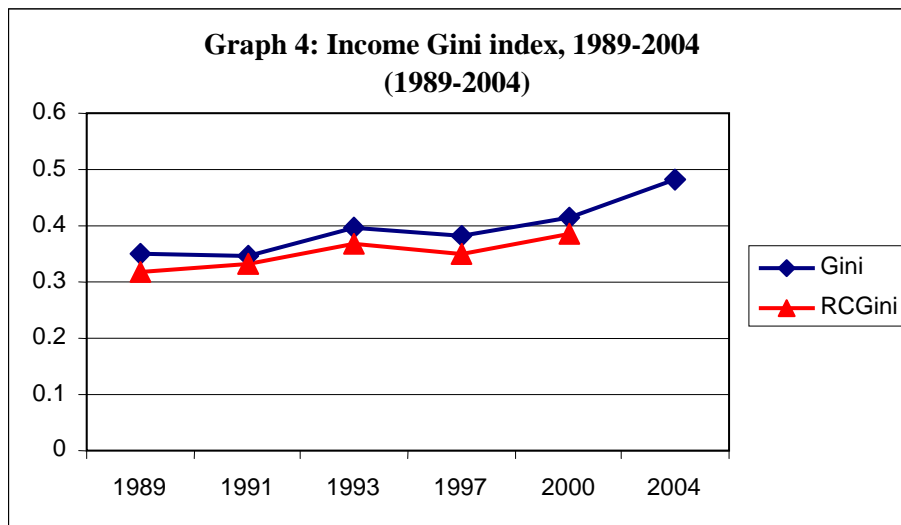
Source: our calculations based on the CHNS data.

In the light of these results, we can underline some important aspects:

- the educational Theil index changed because of a series of forces and variations acting in opposite direction;
- the decrease of educational inequality within the group of “the poor” and the reduction of their population share significantly helped to push down the whole inequality (74.8%);
- educational inequalities within the group of “the not poor” increased, but it didn’t remarkably influence the whole inequality; however, it partially counterbalanced the decrease of the educational Theil index (- 10.4%);
- on the whole, the most part of the Theil index change (64.4%) was generated by within-group dynamics;
- because of the decrease of the ratio $\frac{\overline{HC_1}}{HC}$ for the group of “the poor”, the educational Theil index partially grew, counterbalancing its decrease due to the improvements in educational inequality within this group (- 67.8%);
- on the contrary, the decrease of the ratio $\frac{\overline{HC_2}}{HC}$ pushed down the whole inequality; moreover, this was the most important change influencing the Theil index variation (103.3%);
- on the whole, 35.6% of the educational Theil index change was caused by between-group dynamics.

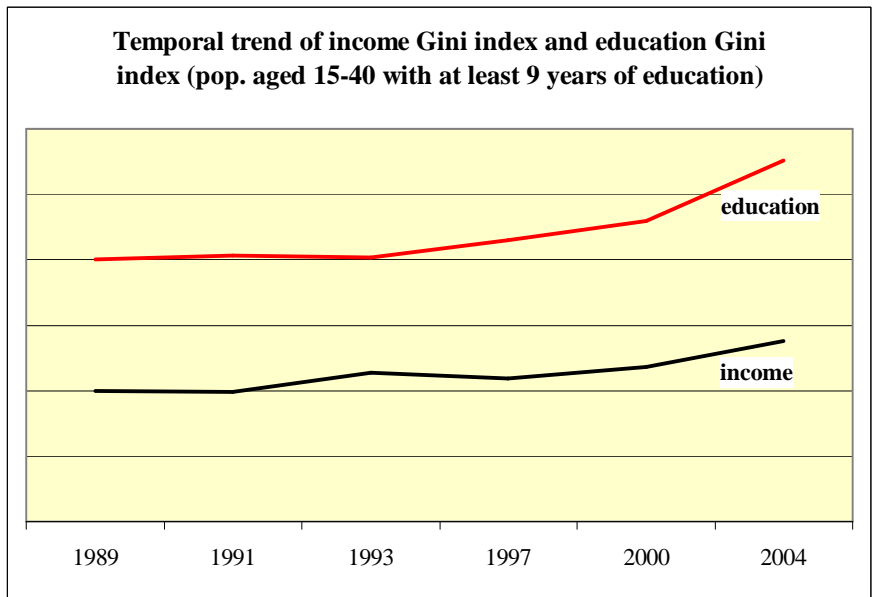
3.2.3 Educational inequality and income inequality: a comparison

When we calculate the Gini index relative to the income distribution, our results confirm those of previous literature: in post-reform China, income inequality constantly increased. In graph 4 we plotted our results with the index Gini values estimated by Chen and Ravallion (2007), based on official statistics; although there is a modest difference in the levels, the two estimations show a trend very similar. The homogeneity between our results and the findings of previous literature makes more reliable our analysis of CHNS data regarding the education distribution.



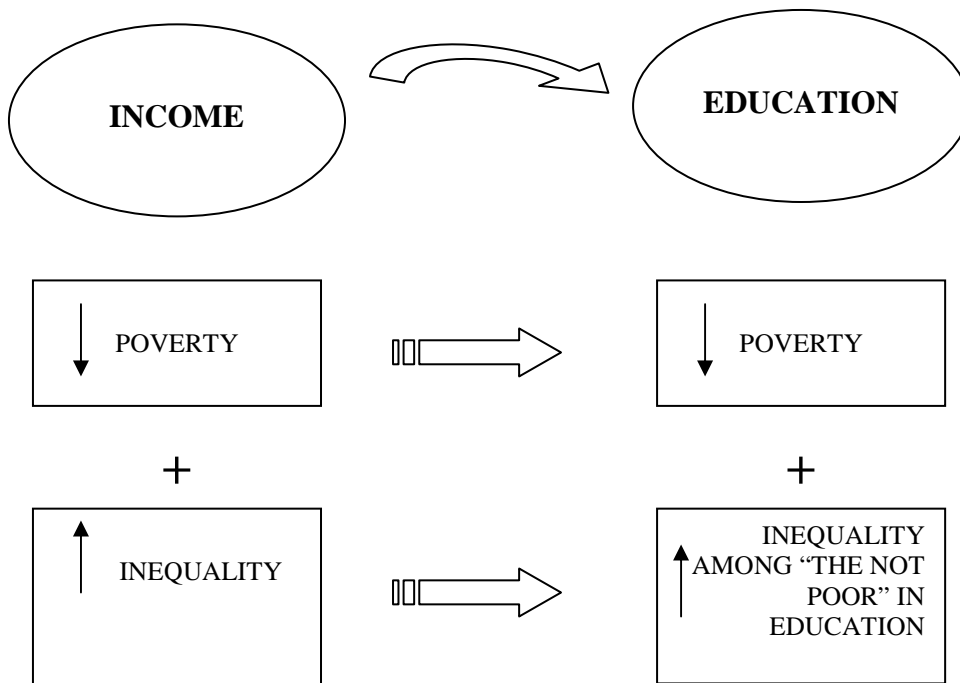
Source: our calculations based on CHNS. The income Gini index is calculated on the whole sample.

If income inequality and educational inequality, calculated on the whole sample, are compared, we observe a contrary movement: the former constantly increased, while the latter followed the opposite trend. At a first look, this phenomenon could appear surprising. However, we saw that educational inequalities raised for the new educated generations, i.e. among people aged 15-40 with at least 9 years of education. It is reasonable suppose that the changes in income distribution were reflected in the opportunities of education of young generations. This point is well described in graph 5. When we compare the trend of income inequality with the evolution of educational inequality among the new generations with at least basic education, we see that they are very similar. How to interpret this finding? The results suggest that in China two different phenomenon occurred (figure 1). On the one hand, the share of people living under the income poverty line decreased, and this caused a parallel reduction of people poor in education which was reflected in a fall of educational inequality indicators; in other words, the reduction of poverty enabled a greater share of people to access to basic education. On the other hand, income disparities rapidly grew, generating an increase of educational inequality among the new generations accessing to medium and high education. This reveal that most young people have concluded the compulsory education (in the 2004, 76%); however, just a little part of these have accessed to the highest levels of schooling.



Source: our calculations based on CHNS. The income Gini index is calculated on the whole sample.

Figure 1: Relationship between income inequality and educational inequality



4. Conclusion.

Although traditional inequality indicators can also be used to analyze educational inequalities, they have to be adapted to some specific features of the education distribution. In fact, in this last case a decrease of the illiteracy rate and educational poverty usually has a strong impact on educational inequality indicators by reducing them. However, this decrease could be misleading if this information is not coupled with a deeper analysis. For this reason, the analysis of educational inequality has to be combined with the study of educational poverty and a series of methodologies making possible to distinguish between distributive improvements due to a decrease of educational poverty and improvements related to a larger diffusion of all the educational levels.

For this purpose, we proposed a decomposition of the educational Theil index and its change by groups: “the poor” and “the not poor” in education. In this way, we were able to distinguish between variations due to what happens below and around the educational poverty threshold, and variations due to dynamics concerning educated people.

By applying this methodology to the Chinese case, the importance of looking beyond the simple and synthetic inequality indicators was confirmed. Over the period 1989-2004, educational inequality indicators constantly decreased not only at national level but also within each province and area. This change was accompanied by a strong reduction of educational poverty and its depth. However, the Theil index decomposition showed that the decrease of educational poverty and the reduction of its depth played a relevant role in improving the whole education distribution. On the contrary, with regard to people accessing to education, educational inequality indicators showed a tendency to increase over the considered period; this reveals that the strongest changes and distributional progress concerned the lowest part of the education distribution, while inequalities among educated people didn't improve. In the light of this we can state that, although the equity in the access to basic education grew, a further effort is necessary to provide a more even access also to the highest levels of education and to reduce polarization between “the poor” and “the not poor” in education.

References

- BARRO R.J., LEE J. (2000), *International Data on Educational Attainment: Updates and Implications*, CID Working Paper No. 42, Harvard University.
- BENJAMIN D., BRANDT L., GILES J., WANG S. (2005), *Income Inequality During China's Economic Transition*, prepared for Brandt L., Rawski T.G. (eds), *China's Great Economic Transformation*, forthcoming volume, Cambridge University Press.

- BENJAMIN D., BRANDT L., GLEWWE P., LI G. (2000), *Markets, Human Capital, and Inequality: Evidence from Rural China*, CHNS paper, UNC, Carolina Population Center.
- CHANG G.H. (2002), *The Cause and Cure of China's Widening Income Disparity*, in "China Economic Review", No.13, p. 335-340.
- CHEN S., RAVALLION M. (2007), *China's (Uneven) Progress Against Poverty*, in "Journal of Development Economics", No.82, p. 1-42.
- CHECCHI D. (1996), *Povert  ed istruzione: alcune riflessioni ed una proposta di indicatori*, Rapporto alla Commissione di Indagine sulla Povert  e l'Emarginazione, Presidenza del Consiglio.
- CHECCHI D. (2001), *Education, Inequality and Income Inequality*, Distributional Analysis Research Programme, Discussion paper No. DARP 52.
- CHECCHI D., GARCIA-PENALOSA C. (2004), *Risk and the Distribution of Human Capital*, in "Economic Letters", Vol.82, Issue 1.
- FANG C., ZHANG X., FAN S. (2002), *Emergence of Urban Poverty and Inequality in China: Evidence from Household Survey*, in "China Economic Review", No.13, p. 430-443.
- FENG J., YU Y. (2006), *Income Inequality and Health in Rural China*, UNU/WIDER Conference "Advancing Health Equity", 29-30 settembre 2006.
- GALBRAITH J.K., KRYTYNSKAIA L., WANG Q. (2004), *The Experience of Rising Inequality in Russia and China During the Transition*, in "The European Journal of Comparative Economics", Vol.1, No.1, p.87-106.
- HANNUM E., WANG M. (2006), *Geography and Educational Inequality in China*, in "China Economic Review", No.17, p. 253-265.
- HECKMAN J.J. (2005), *China's Human Capital Investment*, in "China Economic Review", No.16, p. 50-70.
- HECKMAN J.J., LI X. (2003), *Selection Bias, Comparative Advantage and Heterogeneous Returns to Education: Evidence from China in 2000*, IFAU - Institute for Labour Market Policy Evaluation, Working Paper 2003:17.
- HUANG JRT., KUO C., KAO A. (2003), *The Inequality of Regional Economic Development in China Between 1991 and 2001*, in "Journal of Chinese Economic and Business Studies", Vol.1, No.3, p. 273-285.
- LEE J. (2000), *Changes in the Source of China's Regional Inequality*, in "China Economic Review", No.11, p. 232-245.

- LI H. (2003), *Economic Transition and Returns to Education in China*, in “Economics of Education Review”, No. 22, p. 317-328.
- LIM A.S.K, TANG K.K. (2006), *Human Inequality, Human Capital Inequality and the Kuznets Curve*, CAMA Working Paper n. 8, Australian National University.
- LIN C.A. (2007), *Education Expansion, Educational Inequality, and Income Inequality: Evidence from Taiwan, 1976-2003*, in “Social Indicators Research”, No.80, p. 601-615.
- LOPEZ R., THOMAS V., WANG Y. (1998), *Addressing the Education Puzzle – The Distribution of Education and Economic Reform*, Policy Research Working Paper No. 2031, World Bank, Washington (D.C.).
- LU D. (2002), *Rural-Urban Income Disparity: Impact of Growth, Allocative Efficiency, and Local Growth Welfare*, in “China Economic Review”, No.13, p. 419-429.
- RAM R. (1990), *Educational Expansion and Schooling Inequality: International Evidence and Some Implications*, in “The Review of Economics and Statistics”, Vol.72, No.2, p. 266-274.
- SICULAR T., XIMING Y., GUSTAFSSON B., SHI L. (2007), *The Urban-Rural Income Gap and Inequality in China*, in “Review of Income and Wealth”, Series 53, No. 1, p. 93-126.
- THOMAS V., WANG Y., FAN X. (2001), *Measuring Education Inequality – Gini Coefficients of Education*, Policy Research Working Paper No. 2525, World Bank, Washington (D.C.).
- THOMAS V., WANG Y., FAN X. (2002), *A New Dataset on Inequality in Education: Gini and Theil Indices of Schooling for 140 Countries, 1960-2000*, mimeo, World Bank, Washington (D.C.).
- 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, Washington (D.C.).
- WEI S., WU Y. (2001), *Globalization and Inequality: Evidence from Within China*, NBER, Working Paper No. 8611, Cambridge (MA).
- WU X., PERLOFF J.M. (2005), *China’s Income Distribution, 1985-2001*, in “The Review of Economics and Statistics”, Vol.87, No.4., p. 763-775.
- WU X., XIE Y. (2002), *Does the Market Pay Off? Earnings Inequality and Returns to Education in Urban China*, William Davidson Working Paper No. 454.

XU L.C., ZOU H. (2000), *Explaining the Changes of Income Distribution in China*, in “China Economic Review”, No.11, p. 149-170.

YAO S., ZHANG Z., HANMER L. (2004), *Growing Inequality and Poverty in China*, in “China Economic Review”, No.15, p.145-163.

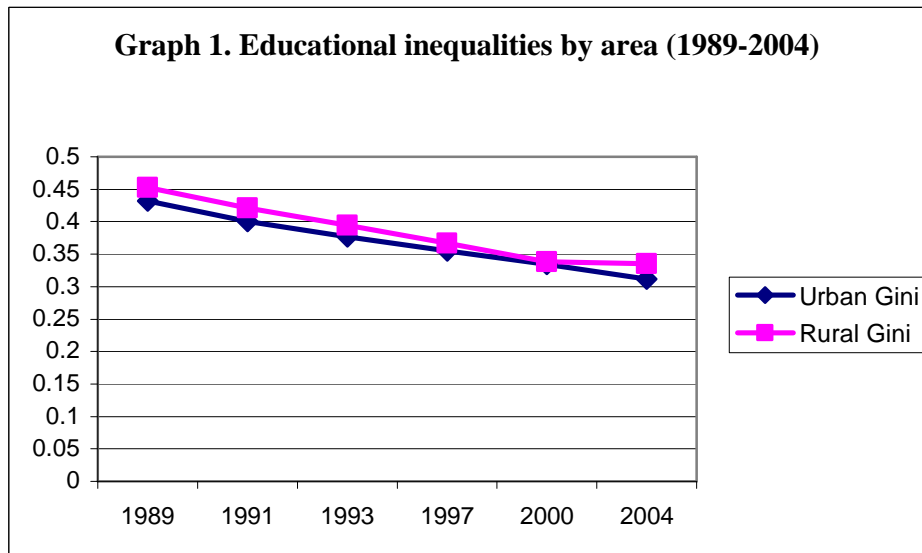
ZHANG X., KANBUR R. (2005), *Spatial Inequality in Education and Health Care in China*, in “China Economic Review”, Vol. 16(2), p. 189-204.

APPENDIX

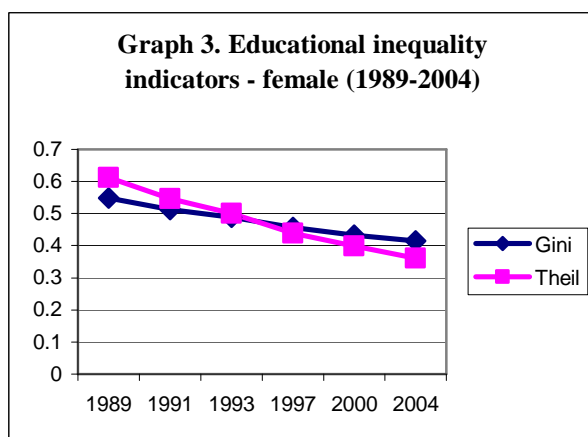
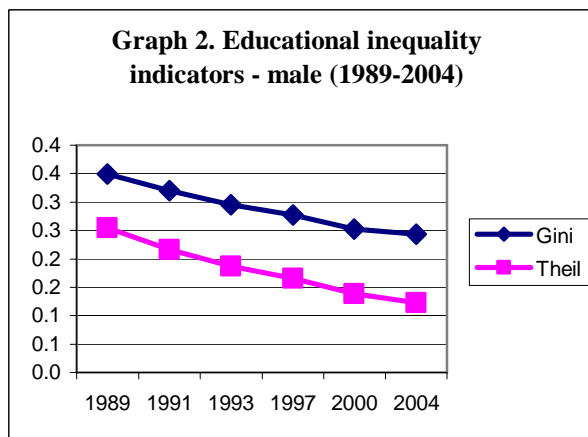
Table 1: Education Gini coefficients by province (1989-2004).

	1989	1991	1993	1997	2000	2004
Jiangsu	0.4556	0.4292	0.4062	0.3703	0.3467	0.3503
Shandong	0.4784	0.4467	0.4146	0.3973	0.3727	0.3592
Henan	0.4574	0.4284	0.4103	0.3843	0.3561	0.3261
Hubei	0.4493	0.4227	0.3984	0.3688	0.3441	0.3454
Hunan	0.3962	0.3582	0.3389	0.3145	0.2887	0.2640
Guangxi	0.3790	0.3493	0.3307	0.3068	0.2823	0.2709
Guizhou	0.5264	0.4802	0.4414	0.4250	0.3991	0.4069
<i>Standard Deviation</i>	0.0494	0.0470	0.0410	0.0428	0.0425	0.0504

Source: our calculations based on the CHNS data.



Source: our calculations based on the CHNS data.



Source: our calculations based on the CHNS data.

Table 2. Average years of education by province.

	1997	2004
NORD	7.98	9.03
Beijing	9.50	10.56
Tianjin	8.38	9.64
Hebei	7.17	8.38
Shanxi	7.68	8.38
Inner Mongolia	7.18	8.17
NORD-EST	8.00	8.71
Liaoning	8.10	8.84
Jilin	8.03	8.80
Heilongjiang	7.86	8.49
EST	7.06	8.11
Shanghai	8.89	10.11
Jiangsu	6.91	7.81
Zhejiang	6.81	7.95
Anhui	6.56	7.49
Fujian	6.73	7.49

Jiangxi	7.05	7.98
Shandong	6.50	7.94
CENTRO	7.15	8.17
Henan	7.10	8.22
Hubei	7.22	8.10
Hunan	7.22	8.16
Guangdong	7.50	8.13
Guangxi	6.61	8.02
Hainan	7.21	8.41
SUD-OVEST	6.20	7.12
Chongqing	6.60	7.25
Sichuan	6.57	7.45
Guizhou	5.85	6.98
Yunnan	5.79	6.82
NORD-OVEST	5.89	7.15
Tibet	3.50	4.40
Shaanxi	7.07	8.26
Gansu	6.13	7.24
Qinghai	4.69	6.80
Ningxia	6.45	7.70
Xinjiang	7.51	8.49
<i>STANDARD DEVIATION</i>	<i>1.1311</i>	<i>1.0823</i>

Source: our calculations based on the CHNS data.

Table 3. Decomposition of educational inequality by provinces, areas and genders (1989-2004, CHNS).

	1989	1991	1993	1997	2000	2004
PROVINCES						
Between	0.8325	0.7985	0.6423	0.8729	1.0453	1.3687
Within	99.1675	99.2015	99.3577	99.1271	98.9547	98.6313
Total	100.00	100.00	100.00	100.00	100.00	100.00
AREAS						
Between	1.0358	1.1152	1.0344	1.2685	1.1541	2.2324
Within	98.9642	98.8848	98.9656	98.7315	98.8459	97.7676
Total	100.00	100.00	100.00	100.00	100.00	100.00
GENDERS						
Between	4.6169	4.5934	4.9971	4.5076	4.7445	4.6610
Within	95.3831	95.4066	95.0029	95.4924	95.2555	95.3390
Total	100.00	100.00	100.00	100.00	100.00	100.00

Source: our calculations based on the CHNS data.

Table 4. Decomposition of educational inequality by provinces and genders (1997 and 2004, CSY).

	1997	2004
PROVINCES		
Between	2.3494	2.2168
Within	97.6506	97.7832
Total	100.00	100.00
GENDERS		
Between	2.0855	1.3512
Within	97.9145	98.6488
Total	100.00	100.00

Source: our calculations based on the CHNS data.