International Journal of Educational Development 37 (2014) 2–10

Contents lists available at ScienceDirect

International Journal of Educational Development

journal homepage: www.elsevier.com/locate/ijedudev

An analysis of education inequality in China

Jun Yang [a], Xiao Huang [b], Xin Liu [a,*]

[a] School of Economics and Business Administration, Chongqing University, Chongqing, China
[b] School of Economics, Chongqing Technology and Business University, Chongqing, China

ARTICLE INFO

Keywords:
Education inequality
China
Gini coefficient
Shapley decomposition

ABSTRACT

This article analyzes both the current situation regarding education inequality in China, and its formation mechanisms. Policies promoting education have lead to remarkable progress in educational attainment, and also effectively decreased educational inequality. However, substantial inequalities in educational attainment remain, even though sustainable progress has been realized. Decomposition results using the Gini coefficient and Shapley value approach based on regression analysis indicate that the greatest contributing factors to educational inequality involve the urban–rural and social stratification divisions. Moreover, the household register system which divides city and country, as well as increasing income inequality is deepening institutional barriers and stratum differentiation. Though gender and regional gaps have been reduced significantly, the population residing in economically disadvantaged areas, especially females, still warrants social concern. In addition, age related decomposition results indicate that increasing educational attainment for the young plays a key role in reducing education inequality. At last, we argue that more educational investment should be allocated to disadvantaged groups and lower income groups; especially eliminating some institutional barriers such as the hukou system, unequal distribution of good quality educational resources, and so on.

1. Introduction

After 30 years of societal reform and an opening up policy initiated by Deng Xiaoping, China has made remarkable achievements in economic growth and education. China's gross national product per capital has reached 2980 US dollars (World Bank, 2010). At the same time, the enrollment ratio of junior middle school students has been maintained at a high level (nearly 100%), which means that China has made nine-year compulsory education virtually universal. Further, more people are able to receive a secondary education, and higher education is popular. As a result, in the area of education, China has made huge progress, and the average years of schooling (AYS) has a sharply increased from 6.794 years in 1996 to 8.28 years in 2008.[1]

Even though educational gaps between various groups still exist, in recent years, increasing attention has been paid to equality issues in education. According to a survey by the China Youth Daily in 2009, only 11.2% of respondents argued that educational gaps were narrowing; school selecting policies, educational gaps between rural and urban areas, and other irrational polices were recognized as the source of education inequality.[2] Actually, as a result of divisive economic structure separating urban and rural areas, increasing income inequality, unbalanced development among different regions, and other inadequate distribution of educational investment and resources, not everyone with normal abilities can acquire the same education (Fig. 1).

Moreover, as most educational resources are controlled by the government, so the government's aims are crucial to education policy and distribution. Despite the fact that education development is viewed as a basic state policy in China, the ratio of public expenditure on education does not keep pace with the GDP growth rate. Fig. 2 tells us that the percentage of government expenditures

* Corresponding author at: School of Economics and Business Administration, Chongqing University, Chongqing 400030, China. Tel.: +86 138 8383 5370.
E-mail address: liuxin@cqu.edu.cn (X. Liu).

[1] Data originates from Table 1.

[2] This survey was carried out by the center of social surveys at the China Youth Daily. It contained 2952 observations across 30 provinces.

http://dx.doi.org/10.1016/j.ijedudev.2014.03.002
0738-0593/

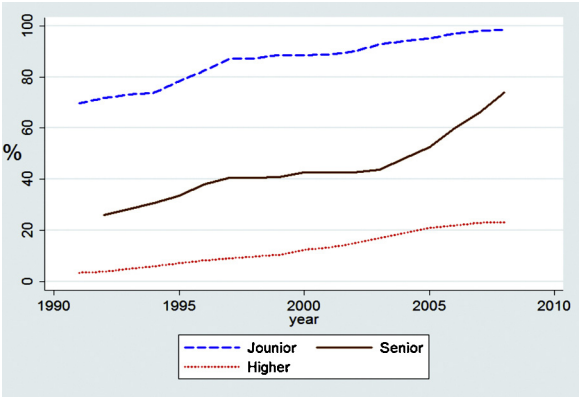

Fig. 1. Chinese education enrollment rate by years. *Note*: Senior School includes regular secondary schools and vocational secondary schools.
Source: China educational statistical year book 2008.

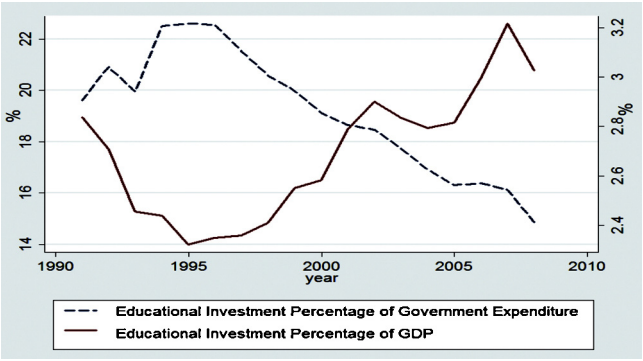

Fig. 2. China's educational investment percentage of government expenditure (GE) and GDP.
Source: China Compendium of Statisticals 1949–2008.

on education still remains low (lower than 3% in most years).[3] Comparing to most OECD countries in 2008, they spent 6.1% of their collective GDP on educational institutions, only nine of 36 countries for which data are available spend 5.0% of GDP or less; and between 2000 and 2008, expenditure for all levels of education combined increased at a faster rate than GDP among most members.[3] It also reveals that only a modest part of national finance revenue is paid to education in China. In fact, insufficient educational investment always leads to unbalanced educational development and education inequality.

In recent years, more and more literature has tried to explain the reason behind China's education inequality. Hannum (1999) summarized the political change in China and drew a comparison between urban and rural areas from 1949 to 1990, the main finding was children in rural are lack of education according to the children in urban. Qian and Smyth (2005) adopted Gini coefficient decomposition to study the educational gaps between rural and urban areas, and also the coastal and inland regions of China. The main finding was that disparities in access to education between rural and urban areas rather than between coastal and inland provinces are the major cause of educational inequality in China. Further, Hannum and Wang (2006) analyzed the Chinese population census for the year 2000, and their results argued that geographic disparity has lead to educational stratification in recent decades.

This paper aims to study the presentation of China's education inequality and its decomposition results, further analyze the reasons and determine what measures should be taken from a public governance view. Of course, we must draw a clear definition that the education attainment and its distribution are confined to national education, some special abilities or talents which should be acquired through apprenticeship training belongs to another important issues.[4]

In this paper, firstly, a scientific and proper measurement of education inequality must be adopted. Gini coefficient has been widely used to study income inequality and could be used to measure education inequality as well, because the education Gini coefficient can effectively represent the change in educational distribution. After measuring education inequality using the education Gini coefficient, a decomposition method based on Gini coefficient will be used to study within-group and between-group contributions to education inequality, according to educational gaps among regions, by gender, between urban and rural areas, and also among different social groups. Based on the empirical findings above, we shall draw a detailed analysis from the point of the educational system and other social factors. At last, Shapley decomposition based on regression analysis will be adopted to study which kind of educational gap contributes to total education inequality most, so that adequate measures can be taken to reduce education inequality.

This paper argues that both national and provincial education inequality is lower than before, and that educational expansion has reduced education inequality significantly. The urban–rural division and social stratification division are the greatest contributors to education inequality. Moreover, the household register system dividing city and country, and increasing income inequality are deepening institutional barriers and stratum differentiation. According to our investigation, although gender and regional gaps have been reduced significantly, the population from poorer areas (especially for females) still merits social concern. In addition, through decomposing age, we also find that the overall education inequality drops sharply as age decreases, which is mainly a product of higher educational attainment among the young.

2. The extent of China's education inequality

To what extent can we measure education inequality? The standard deviation of years of schooling is often chosen as a measure of education inequality in a few studies,[5] but such a method only measures the dispersion of schooling distribution in absolute terms. To measure the relative inequality of the schooling distribution, developing an indicator for education Gini coefficient is advisable, so we have adopted an indirect method originated by Tomas et al. (2003) to calculate education Gini coefficient based on educational attainment data. The education Gini formula is shown in Eq. (1).

$$E_L = \left(\frac{1}{\mu}\right)\sum_{i=2}^{n}\sum_{j=1}^{i-1} p_i|y_i - y_j|\,p_j \tag{1}$$

where E_L is the education Gini based on educational attainment distribution, large population; μ is the average years of schooling for the concerned population; p_i and p_j stand for the proportions of population with certain levels of schooling; y_i and y_j are the years

[3] *Data resource*: OECD (2011), Education at a Glance 2011: OECD Indicators, OECD Publishing.

[4] For example: If someone is talented in making Sushi, he should have become an apprentice to a famous Sushi chef just after graduation from junior high school. Acquiring higher educational achievement in national education system may not be his best choice. So in this situation, his own decision does no matter to education inequality.

[5] For example: Ram (1990).

Table 1
Education Gini coefficients and average years of schooling.

Region	Year							
	1996		2000		2004		2008	
Index:	Gini	AYS	Gini	AYS	Gini	AYS	Gini	AYS
All Nation	0.2829	6.79	0.2377	7.62	0.2385	8.01	0.2255	8.27
Beijing	0.2269	9.58	0.2109	9.99	0.2061	10.56	0.1977	10.97
Tianjin	0.2446	8.02	0.2209	8.99	0.2132	9.64	0.2048	9.88
Hebei	0.2554	6.89	0.2162	7.74	0.2043	8.38	0.1875	8.36
Shanxi	0.226	7.56	0.2047	8.02	0.1882	8.38	0.1866	8.81
Inner Mongolia	0.2844	7.12	0.2488	7.76	0.2431	8.17	0.228	8.37
Liaoning	0.2208	7.85	0.2093	8.41	0.1911	8.84	0.1991	9.08
Jilin	0.2364	7.76	0.214	8.24	0.1916	8.80	0.1958	8.89
Helongjiang	0.2415	7.73	0.2096	8.25	0.1829	8.49	0.1866	8.70
Shanghai	0.2373	9.01	0.2212	9.30	0.2174	10.11	0.2023	10.55
Jiangsu	0.2883	7.02	0.2329	7.85	0.2552	7.81	0.2254	8.44
Zhejiang	0.2763	6.71	0.2449	7.46	0.2734	7.95	0.2531	8.24
Anhui	0.2903	6.30	0.2632	6.98	0.2625	7.49	0.2618	7.44
Fujian	0.2973	6.29	0.2357	7.49	0.278	7.49	0.2512	7.80
Jiangxi	0.2579	6.54	0.2175	7.55	0.2263	7.98	0.2206	8.26
Shandong	0.3148	6.42	0.2438	7.58	0.2507	7.94	0.2178	8.28
Henan	0.2622	6.85	0.2145	7.72	0.2043	8.22	0.2013	8.34
Hubei	0.2768	6.94	0.2348	7.77	0.2441	8.10	0.2316	8.49
Hunan	0.2509	6.94	0.2049	7.80	0.2145	8.16	0.2115	8.43
Guangdong	0.2449	6.79	0.2054	8.07	0.2138	8.13	0.1943	8.77
Guangxi	0.2364	6.70	0.203	7.57	0.2209	8.02	0.1909	7.98
Hainan	0.2889	6.64	0.2374	7.68	0.2108	8.41	0.2171	8.35
Chongqing			0.2301	7.28	0.2496	7.25	0.2174	7.79
Sichuan	0.2902	6.43	0.2422	7.06	0.2422	7.45	0.2413	7.51
Guizhou	0.3739	5.69	0.3086	6.15	0.2838	6.98	0.2604	7.05
Yunnan	0.359	5.62	0.2946	6.33	0.2805	6.82	0.2565	6.90
Tibet	0.5994	2.92	0.5946	3.43	0.4437	4.40	0.4271	4.71
Shaanxi	0.3087	6.81	0.2454	7.71	0.2433	8.26	0.2343	8.51
Gansu	0.4044	5.73	0.3273	6.54	0.3111	7.24	0.2963	7.17
Qinghai	0.4935	4.90	0.3954	6.12	0.3334	6.80	0.3072	7.26
Ningxia	0.3467	6.50	0.309	7.03	0.286	7.70	0.2452	8.13
Xinjiang	0.2851	7.21	0.2486	7.73	0.2354	8.49	0.2157	8.56

Note: (1) Gini and AYS are short for education Gini coefficient and average years of schooling respectively. (2) Chongqing's index in 1996 is absent because it has been considered a municipality since 1997; before that time it was not categorized as a single province.

of schooling at different educational attainment levels; n is the number of levels in the attainment data.

Barro and Lee (1993) divided the population into seven categories ($n = 7$) including: no-schooling, partial primary, complete primary, partial secondary, complete secondary, partial tertiary and complete tertiary. However, China maintains educational statistics using its own structure. It divides the population aged 6 and over into five groups ($n = 5$), which includes: no schooling, primary school, junior secondary school, senior secondary school, college and higher level, and the corresponding education attainment years are 0 year, 6 years, 9 years, 12 years, and 16 years, respectively. In fact, we can calculate national and provincial educational inequality from 1996 to 2008 in China, and the related original survey statistics come from the China Statistical Year Book 1997–2009.[6] For simplicity, we only show some representative years of education inequality (see Table 1).

From Table 1, we can see China's education Gini coefficient has been reduced considerably since 1996, no matter the whole country or individual provinces. But there are still some phenomena we should pay attention to. For one thing, most western provinces' education Gini coefficients are higher than that of eastern provinces. For example, Beijing's education Gini coefficient in 2000 and 2008 are 0.2377 and 0.2255 respectively, but the corresponding numbers of Guizhou province are 0.3086 and 0.2604. This indicates that education inequality in the western regions is usually higher than in eastern regions, which mirrors the same economic development pattern between regions.[7] For another thing, if one's AYS index is higher, its education Gini coefficient is lower accordingly. For example, the AYS indices of Beijing, Tianjin, Shanghai have maintained a higher level in recent years, and their education Gini indices are lower correspondingly. At last, for any unit the deeper the initial education inequality is, the faster the reduction of educational gaps will be. During the past 13 years from 1996 to 2008, education inequality in Tibet, Qinghai and Gansu has decreased 28.87%, 37.5%, and 26.73%, respectively, but such indices for Beijing, Shanghai and Tianjin are only 20.2%, 14.74%, and 16.27%. To surmise, we can conclude that each province's education inequality is lower than before, but gaps between different regions still exist.

3. Decomposing education inequality using Gini coefficient

3.1. Data resource

With the basis of measuring China's national and provincial education inequality, we draw a general acknowledgment. Such empirical findings can only describe the extent of educational inequality and its tendencies, but not sufficient for analyzing the

[6] China's education attainment survey data was first published in 1996 by the National Bureau of Statistics of China, so the education Gini coefficient before 1996 cannot be calculated.

[7] As a result of unbalanced development among different areas, 31 provinces in mainland China have been divided into three regions. The eastern region is usually recognized as relatively-developed, which includes Beijing, Tianjin, Hebei, Liaoning, Shanghai, Jiangsu, Zhejiang, Fujian, Shandong, and Guangdong provinces. The central region's development is less than the Eastern region but greater than the Western region. The central region includes: Shanxi, Hunan, Hubei, Hunan, Jiangxi, Anhui, Jilin, and Heilongjiang. At last, the rest of the provinces belong to the Western region, which is the least developed of the three.

Table 2
Variable Summary.

Variable	Description of variable
province	28 provinces are included in the survey and divided into three parts according to their GDP per capital. Eastern regions: Beijing, Tianjin, Hebei, Liaoning, Shanghai, Jiangsu, Zhejiang, Fujian, Shandong, and Guangdong. Central regions: Jilin, Heilongjiang, Anhui, Jiangxi, Henan, Hubei, Hunan, and Shanxi. Western regions: Inner Mongolia, Guangxi, Hainan, Chongqing, Sichuan, Guizhou, Yunnan, Shannxi, Gansu, and Xinjiang.
gender	Dummy variable, 1 = male and 0 = female.
hukou	Dummy variable, 1 = agricultural and 0 = non-agricultural. Registered permanent residence system in China differentiates the agricultural and non-agricultural registered permanent residence strictly, resulting in the forming of intersected structure of town and country.
age	Respondents' age in 2005.
edu	Years of schooling attainted in 2005.
income	All household income last year, measured by thousands of yuan in RMB.

mechanism behind the truth. As a result, here we adopt the China General Social Survey (CGSS) from 2006 and a decomposition method using Gini coefficient to study why education inequality has formed.

A summary of the data is necessary as a starting point for further analysis. CGSS was organized by the department of sociology of Renmin University of China and the division of social science of the Hong Kong University of Science and Technology. The CGSS is a national survey comprising 28 provinces and 10,000 observations of individuals between 18 years old and 69 years old, so analysis with the basis of this data is representative and significant. After analyzing the samples carefully, some useful information was chosen, such as an individual's educational attainment, income and family characteristics. However, dropping some observations were unavoidable, as missing values existed. Here in Tables 2 and 3 we describe summary statistics of key variables used in our study. Note that the percentage of female and agricultural observations hovers at nearly 50% of all samples, and values of income, education and age are nearly normally distributed.

3.2. Method

Generally speaking, the decomposition of Gini coefficient is initially used to analyze income inequality and then subsequently expanded to study education inequality. Gini coefficient decomposition includes two main approaches: one is decomposition by subgroup and calculation between-group contribution and within-group contribution, such as population between rural and urban, male and female. Another analytical method is decomposition by components. For example, total income can be expressed as different resources so that each component's contribution can be calculated. When studying education inequality, the first approach is usually used.

Table 3
Sample description.

7752 observations in all							
Gender = 1; 3799 obs (49%)					Gender = 0; 3953 obs (51%)		
Hukou = 1; 3724 obs (48%)					Hukou = 0; 4028 obs (52%)		
Percentiles	5%	10%	25%	50%	75%	90%	95%
edu	3	5	6	9	12	14	15
income	3	4.1	8	15	28	48	60
age	21	24	32	41	52	60	64

In early studies, Zhang and Li (2002) proposed a formula of Gini decomposition, that is:

$$G = P_1^2\left(\frac{\mu_1}{\mu}\right)G_1 + P_2^2\left(\frac{\mu_2}{\mu}\right)G_2 + G_B \tag{2}$$

where P_i, μ_i, G_i $(i = 1, 2)$ represent the proportion of the population, mean education, and Gini coefficient for two exclusive subgroups of the population. G_B is the between-group contribution to the total.

However, the decomposition method above has two limits. First, it only describes the definition of decomposing into two subgroups. Second and most significant, if there is no overlapping data among each subgroup, Gini coefficient would be additive and totally decomposed, otherwise, a biased result would be forthcoming and a residual term R would likely be left (Ebert, 1988). Unfortunately, such a perfect situation does not appear in most cases. As we know, Gini decomposition by subgroups originated from Soltow (1960) and developed by Pyatt (1976) and Mookherjee and Shorrocks (1982). For the existence of the residual term R, many studies have tried to interpret it. For example, some argue R is an interaction effect resulting from overlapping between each subgroup (Silber, 1989), while others explain R as stratification (Yitzhaki and Lerman, 1991).

Lambert and Aronson (1993) provide a solid understanding of the residual term R as a sub-area of the Lorenz diagram, so that implications for inequality decomposition analysis are briefly considered. Moreover, such an approach is not the only form of Gini decomposition on offer, many other specifications have been suggested (Shorrocks and Wan, 2005). In order to avoid the existence of overlapping data and a biased result, we shall adopt the approach mentioned in this paragraph to decompose the educational Gini coefficient.

Let G be the Gini coefficient and let the population subgroups be indexed by $k = 1, 2, \ldots, m$.[8] The decomposition takes the form:

$$\begin{aligned} G = G(y^1, y^2, \ldots, y^m) &= \frac{2}{n^2\mu}\sum_{k=1}^{m}\sum_{i\in N_k} r_i(y_i - \mu) \\ &= \frac{2}{n^2\mu}\sum_{k=1}^{m}\left\{\sum_{i\in N_k} i(y_i - \mu_k) + \sum_{i\in N_k} i(\mu_k - \mu) + \sum_{i\in N_k}(r_i - i)y_i\right\} \\ &= W + B + R \end{aligned} \tag{3}$$

where

$$W = \frac{2}{n^2\mu}\sum_{k=1}^{m}\sum_{i\in N_k} i(y_i - \mu_k) = \sum_{k=1}^{m} v_k^2 b_k G(y^k) \tag{4}$$

$$\begin{aligned} B &= \frac{2}{n^2\mu}\sum_{k=1}^{m}\sum_{i\in N_k} i(\mu_k - \mu) = \sum_{k=1}^{m} b_k v_k\left[\sum_{j=1}^{k} v_j - \sum_{j=k}^{m} v_j\right] \\ &= G(y^{\bar{1}}, \ldots, y^{\bar{m}}) \end{aligned} \tag{5}$$

The first term of the right-hand side in Eq. (3) is the within-group contribution, while the second term is the between-group component of education inequality. R is a residual, which is zero if the subgroup observation ranges do not overlap. When we aim to analyze education inequality caused by urban–rural division and gender difference, population should be divided into two subgroups ($m = 2$). In other situations, such as decomposing among various regions, ages and income status, the subgroup numbers are 3, 9 and 5 respectively.

[8] *Note*: Here the Gini coefficient used is different from the formula in Eq. (1); the former adopts survey data to calculate and the latter use tabulate data to estimate.

Table 4
Average years of schooling (AYS) and Gini for each subgroup.

		AYS	Educational Gini coefficient
Total sample		8.877064	0.21911
By hukou	Rural	7.223953	0.17756
	Urban	10.40541	0.21695
By gender	Male	9.092393	0.20959
	Female	8.670124	0.30935
By region	East	9.537973	0.2057
	Central	8.326907	0.21541
	West	8.457442	0.23372
By income rank	Rank = 1	6.663455	0.23261
	Rank = 2	7.721758	0.20753
	Rank = 3	9.044802	0.1877
	Rank = 4	10.19983	0.17674
	Rank = 5	11.3704	0.16259
By age	Age ∈ [18,22]	10.92789	0.1468
	Age ∈ (22,26]	11.13818	0.17173
	Age ∈ (26,32]	9.956243	0.20183
	Age ∈ (32,38]	8.929315	0.21586
	Age ∈ (38,44]	8.431173	0.29018
	Age ∈ (44,50]	8.80884	0.19527
	Age ∈ (50,56]	7.77894	0.22256
	Age ∈ (56,62]	7.39006	0.25063
	Age ∈ (62,70]	7.481262	0.26852

Note: (1) Income rank = 1 indicates the lowest income household (first quintile group); then rank from 2 to 5 represents lower middle income household (second quintile group), middle income household (third quintile group), upper middle income household (fourth quintile group) and highest income household (fifth quintile group) respectively. (2) Other explanations are the same as shown in Table 2.

3.3. Empirical result

Table 4 shows the average years of schooling (AYS) and educational Gini coefficient for each subgroup, which describes the sample structure and helps to understand the decomposition result. Please note the AYS is 8.87 and educational Gini coefficient is 0.21, as compared to the national result for 2004 (as shown in Table 1) of 8.01 and 0.23, respectively. Therefore, the sample we selected is comparable with a survey completed by the National bureau of Statistics of China. That means related empirical results based on CGSS can reflect the general status of China.

Now, we shall analyze the decomposition result and its policy implication cases.

When decomposing for rural and urban subgroups (variable hukou), contribution from the Gini coefficient within subgroups is as good as that from between subgroups. This means overall education inequality depends largely not only on urban–rural differences, but also on variations within rural differences and inner cites.[9] For one thing, the educational disparity between urban and rural has not been eliminated, even as the overall AYS significantly increased and education inequality has had a sharp decline in recent decades. Concerning educational attainment, the AYS in urban and rural are 10.4 years and 7.22 years respectively, so educational gaps between urban and rural still exist. Regarding the distribution of good educational resources, the urban areas are also superior to the rural. From Table 6 we can see that there are always more teachers with a college education in urban than in rural areas. Consequently, students from urban areas can acquire a relatively good-quality education, have better chances of progressing to the next educational level, and demonstrate overall better achievement in education.

Table 5
Decomposition by subgroups.

Subgroups	Within-group contribution (%)	Between-group contribution (%)	*R* (%)	*W*/*B*
By hukou	44.24	40.82	14.94	1.08
By gender	59.38	5.42	63.8	10.95
By region	34.28	15.03	50.69	2.28
By income rank	17.63	48.03	34.34	0.36
By age	12.97	31.96	55.07	0.4

A major consideration is the cause of within-group inequality. As we know, China has made compulsory education of nine years universal during the past 30 years, and it has decreased the educational attainment gap both in urban and rural areas. However, the rising income inequality may reduce the positive effect from education expansion,[10] even leading to more education inequality (Yang et al., 2009). Moreover, as a result of the scarcity of educational resources (especially for good quality educational resources most prevalent in urban areas), such negative effect from income inequality causes advantaged groups to receive a disproportionate amount of the available good quality education. The kind of education inequality originating from social stratification division is becoming a new rising factor, leading to education inequality in the process of urbanization and industrialization.

Gender is usually organized as an important factor in education inequality, and the literature often shows that gender plays a key role in the enrollment rate, educational attainment and so on, such as Barro and Lee (1993) and Zhang and Li (2002). Through observing the decomposition results in Table 5, its apparent that the within-group component is the main contributor to overall education inequality. China's education inequality between genders is at a low level because of education expansion and lasting gender equality promotion. Furthermore, when considering gender disparity in rural areas, the male and female's AYS are 7.53 and 6.9, while such values in urban areas are 10.64 and 10.19.[11] In addition, we also use the formula in Eq. (1) to calculate the AYS and education Gini coefficient for the male and female (see Table 7). Note that the AYS gap declines from 1.33 in 1996 to 0.96 in 2008, while at the same time, education Gini coefficient for the male and female also decreases 15.2% and 23.8% during the process.

Thanks to the compulsory education system and gender equity promotion, the gender gap in educational attainment has been greatly eliminated in the past decades. However, poverty is still an important factor contributing to gender inequality in rural education. In China's poverty-stricken rural areas, as a result of a deficient supply of diversified household livelihood capitals, men enjoy priority over women in receiving education (Dong et al., 2008). In short, though educational disparity between male and female has been sharply decreased, such differences still exist and deserve further attention.

When decomposing for the east, central and west regions, empirical results show the contribution rate from within-group components is double than that of between-group components. Of course, the within-group contribution mainly comes from gaps between the relatively advanced and the comparatively backward provinces in the same region. For example, Tibet and Shaanxi all belong to west region but the former greatly trails the latter in educational development. The implication may lie in that each province has its own characteristics, so more regionally appropri-

[9] Not the same as our finding, Qian and Smyth (2005) have founded total education inequality in 1995 and 2000 mainly comes from urban–rural differences. The reason for non-uniform results might lie in the data and decomposition methods selected. Nevertheless, the educational gap between urban and rural areas is an important factor when considering education inequality in China.

[10] There is no official report about income Gini coefficient in China, but according to world bank's survey, China's income Gini coefficient is 0.415 in 2005 (World Bank, 2010).

[11] *Data resource*: author's calculation.

Table 6
Comparison of teachers between urban and rural.

Year	Percents of primary school teachers' degree above three years college (%)			Percents of junior high school teachers' degree above bachelor (%)			Percents of senior high school teachers' degree above bachelor (%)		
	City	Town	Rural	City	Town	Rural	City	Town	Rural
2003	64.4	49.05	31.76	48.67	22.12	14.27	86.69	70.95	59.45
2004	71.34	58.41	40.14	55.02	28.01	18.99	88.87	75.91	65.36
2005	78	67.17	47.48	61.75	34.5	24.33	89.22	80.49	70.52
2006	85.29	75.86	58.53	71.96	46.79	35.96	94.15	87.2	81.2

Source: Author's calculation from "China's educational year book" 2004–2007.
Note: Because of statistics non-uniformity, urban area is divided into the subcategories of city and town for this analysis.

Table 7
Education expansion and inequality in China from 1996 to 2008.

	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
Average years of schooling above 6 aged (AYS)													
Male	7.45	7.63	7.69	7.77	8.14		8.26	8.42	8.5	8.38	8.54	8.66	8.73
Female	6.12	6.36	6.46	6.56	7.06		7.18	7.37	7.5	7.26	7.52	7.7	7.77
Education Gini coefficient													
Male	0.23	0.226	0.223	0.222	0.202		0.208	0.208	0.205	0.208	0.203	0.198	0.195
Female	0.335	0.321	0.316	0.312	0.273		0.283	0.278	0.271	0.288	0.271	0.262	0.255

Source: Author's calculation, original data are from China Statistic Year Book 1997–2009.
Note: Data in 2000 and 2001 both come from the 5th national census. While others are obtained from the sample survey on population changes each year by National Bureaus of Statistics, the sample fraction keeps nearly 1‰ of all population.

ate policies should be implemented. Specifically, some less developed areas, which have higher education inequality and a lower education expansion level, need more support including educational resources, financial aid and investment. However, between-group disparity still deserves special attention. As a result of unbalanced development between areas, developed areas have more educational funds and higher educational development, which means educational gaps between different areas could not be quickly eliminated. The comparison in Table 8 indicates the share of educational investment among each region. Although the eastern region contains 10 provinces, it has received half of the total educational investment in recent years. The central area with eight provinces accounts for 24.57%. In contrast, the west with 10 provinces only takes one-fifth in all. As Tsang (1996) said, the financial reform of basic education in China has moved rapidly away from a centralized system with a narrow revenue base to a decentralized system with a diversified revenue base, but the current system is marked by notable weakness in terms of glaring inequalities and significant inefficiencies. In fact, China's central government should increase educational transfer payments to less developed provinces, especially for some poverty stricken areas mainly located in the west. In addition, the establishment of a financial system governed by each provincial institution is advisable.

Table 8
Educational investment among regions.

	Educational investment (billion yuan)			Educational investment (%)		
	East	Central	West	East	Central	West
1998	130.6766	67.987	52.3374	52.06	27.08	20.86
1999	154.2288	77.9024	60.1163	52.77	26.65	20.58
2000	176.0917	87.3749	68.3192	53.09	26.33	20.58
2001	206.2421	96.9406	78.1415	54.08	25.42	20.5
2002	290.6097	134.9673	116.8338	53.57	24.88	21.55
2004	333.9299	151.291	129.2877	54.34	24.61	21.05
2005	391.1715	176.1342	149.394	54.57	24.57	20.85
2006	452.4777	205.9642	173.9176	54.36	24.74	20.86
2007	454.0489	215.655	185.135	53.11	25.22	21.67
2008	550.1017	272.0031	241.2853	51.73	25.57	22.6

Resource: China Statistic Year Book 1999–2009.
Note: (1) Regional classification is the same as shown in Table 1. (2) Educational investment = government appropriation for education + funds from private schools + donations and fund-raising for running schools + income from teaching research and other auxiliary activity + other educational funds. (3) Educational investment is measured by RMB.

Next, let us focus on the impact of social stratification division. Individual income is usually recognized as a crucial factor which can influence one's educational attainment, so decomposing educational inequality based on income level helps to analyze the extent to which education inequality can be influenced by social stratification division. The empirical result in Table 5 shows that the between-group component contributes more than the within-group component. From the point of education attainment, the AYS of the first quintile group is 6.66 years, while that of the fifth quintile group is 10.39 years (see Table 4). This reveals a clear implication that income influences one's education attainment; the higher income group you belong to, the more education you may acquire. Note that of the many factors affecting income inequality, the urban–rural income gap is recognized as the most important one.[12] As there is still 54.32% of the population which lives in rural areas, their lower income level inhibits them from receiving as good a quality of education as the urban population (Table 9).

Further, when investigating a respondents' father's work unit in each income level, we can see people whose father worked as a peasant comprises the majority of the lowest 20% income group (first quintile), while offspring from the state-owned sector takes at least 40% of the fifth quintile group. The results demonstrate a tendency that if one's father works in a state-owned sector he would have more opportunity to enter the high income group, which is always accompanied by higher educational attainment. Actually, the father's work unit represents the inequality of educational opportunity and resources, and such gaps come from family background and social power which not only denominates one's educational achievement but also affects one's work and income in the long term. As a result, social stratification division has an important impact on one's educational attainment, and the division between urban and rural makes such an effect deeper.

Our findings support empirical evidence for the education and sociology theory. It argues that early educational development

[12] The ratio of annual per capital disposable income of urban to per capital net income of rural households stands at 2.57 in 1978, while such value in 2008 raises to 3.31.

Table 9
Respondents' father's work unit.

	State-owned sector (%)	Non-state sector (%)	Peasant (%)
First quintile group	10.34	6.22	83.44
Second quintile group	21.57	11.04	67.39
Third quintile group	34.68	14	51.32
Fourth quintile group	31.02	13.1	55.88
Fifth quintile group	41.87	20.16	37.97

Source: Author's calculation from CGSS data.
Note: (1) It reveals respondents' father's occupational category when respondent was 18 years old. China has been experiencing fast development with a distinct character of a state-owned sector strongly influencing economics and society. Here, the state-owned sector contains people who belong to Party, government organizations, and enterprises owned by the nation; non-state sector includes self-employed, private enterprises from the China mainland, Taiwan, Hong Kong, Macao and foreign direct investments. The rest are works as peasants. (2) Income group is the same as shown in Table 2.

cannot reduce education inequality based on family background, because the new rising educational opportunities are available for all students but not for the disadvantaged group only. So distribution of educational opportunity based on family background may change if the institution of educational resources has not changed. This is called the theory of keeping maximum inequality. The theory argues that education has the effect of promoting equality, but such an effect would happen only if enough education is gained by the advantaged group. After that, education inequality based on family background would decrease. In China, although the government has taken some measures to reduce education inequality, such as scholarship, financial aid, and transfer payments to less developed areas, education inequality based on family background has not yet been solved fundamentally. For example, because of the rural–urban divide in the household registration system, students from rural areas cannot move to cities to acquire good-quality education, despite the fact that most of their parents are doing business and seeking jobs in cities. Apart from this, as a result of unequal distribution and scarcity of good quality education, students who want to get a good education but are not qualified for the exam must pay extra expenses when selecting a school. According to our survey, such fees are approximately 35,000 RMB,[13] while annual per capital disposable income of urban household and annual per capital net income of rural households are only 15,781 yuan and 4761 yuan respectively. In fact, expenses for selecting a school place a heavy burden on a normal household. In short, with the background of income inequality, social stratification division deepens educational gaps between groups.

At last, we focus on the decomposition results based on decomposing age, and we can see that the between-group component contributes more than the within-group component. Moreover, the younger group tends to have a higher AYS index and lower Gini coefficient, while the older group is in the opposite situation. Why leads to such a result? Maybe the social and political conditions in the 1970s forced some people in their around 50/60s now to give up to go to college. This generation seems to cause a downward bias in the disparity in years of schooling in China. At another point, with the background of education expansion in last few decades, the decomposition exercise reveals an interesting pattern that the education gap within each group plays an important role in reducing education inequality.[14] Though educational gaps among different age groups may last a short time, the increase of education attainment for the young helps to lower total education inequality. The meaningful implication is that education inequality can be reduced further if more attention will be paid to the young, especially regarding equal opportunity for school entrance and good quality education.

4. Which contributes to education inequality most?

The above analysis has studied within-group and between-group contributions in the view of urban–rural division, regional gaps, gender disparity and social stratification division. Comparing all these factors above, which contributes to overall education inequality most? Answering this question can help us to understand the cause which leads to education inequality and its implication. Decomposition based on regression analysis can be used to do this study: it not only allows all factors which influence inequality to be identified and quantified, but also includes a significant number of variables, while other traditional decomposition methods such as Gini and General Entropy do not.

In comparison to other approaches,[15] the method proposed by Wan (2002, 2004) combines regression and Shapley decomposition theory so three advantages are available: First, it has no limitation on the measurement of inequality, no matter the Gini coefficient or General Entropy; Second, it allow more variables; Third, it has no limit on the syntax of the regression model, both dummy variables and interaction terms are permitted; even one equation of simultaneous equation model (SEM) is applicable.

The first step is to apply Shapley value decomposition to estimate a function of educational attainment. As we know, there are many factors affecting education, and we intend to select some significant variables according to the CGSS data. The function we shall estimate expresses as follows:

$$\log(\text{edu}) = F(\text{gender}, \text{hukou}, \text{age}, \text{income}, \text{east}, \text{central}) \quad (6)$$

Here, edu and income represent years of schooling and household income respectively. Gender, hukou, east, central are dummy variables; all their definitions are shown in Table 2. Because the variable edu is approximately normally distributed, so Eq. (6) can be set as a semi-log model. In addition, heteroscedasticity is a common problem in a cross-section data pattern, which can cause the standard error of the estimated coefficient to be ineffective. In order to test whether heteroscedasticity exists, we adopt two testing methods to make the result robust. From Table 10 we can see, no matter what testing approach is used, heteroscedasticity exists when using OLS. As a result, another method of FGLS is suitable to be used.

From Table 11, we can see that all dependent variables are significant and this result can be used to do Shapley value decomposition effectively. In addition, different inequality indices may lead to a non-uniform result, so we adopt three approaches together to ensure the robustness of the decomposition result.[16]

According to the decomposition results in Table 12, no matter which inequality measures are used, the division between urban and rural contributes to total education inequality the most. It is not hard to understand that the disparity between urban and rural is still the primary part of education inequality. As a result of the disparity in resource allocation between urban and rural, over a long period of time, the urban area achieves greater development than the rural area. In fact, the largest part of educational investment is present in cities so that teachers with higher degrees and modern equipment for teaching are common in cites, while

[13] Maybe such fees are extra higher in some big cities.
[14] Lin and Yang (2009) exert an empirical study on education inequality in Taiwan. The results show that overall education inequality drops sharply, which is mainly contributed within the age group component after decomposing education inequality using the decomposable Theil index.
[15] Such as some earlier methods developed by Blinder (1973) and Oaxaca (1993).
[16] Three inequality indexes are Gini coefficient, GE(0) and GE(1). GE is short for General Entropy, when the parameter aOaxaca= 1, it is always called Theil index; or else called second Theil index when $a = 0$.

Table 10
Heteroscedasticity test.

B–P test (Breusch and Pagan, 1979)	White test (White, 1980)
808.17 (0.00)***	399.84 (0.00)***

Note: (1) Both B–P test and White test's null hypothesis is homoroscedaticity. (2) Numbers in the parenthesis are *P*-value.
*** Significant at 1% level.

Table 11
FGLS regression result.

Dependent variable: log (edu)			
Independent variable	Coefficient	*T*-Statistics	Significant level
gender	0.0758	10.32	***
hukou	−0.3126	−38.88	***
age	−0.01	−32.44	***
income	0.1152	26.04	***
east	0.0354	3.71	***
central	0.0168	1.64	*
_constant	2.2934	105.86	***
Adj-R^2 = 0.4358		*F*-Statistic = 997.19***	

* Significant at 10% level.
*** Significant at 1% level.

Table 12
Shapley decomposition result.

Factors	Education inequality index used		
	Gini	GE(0)	GE(1)
Gender	3.88	1.03	1.18
Hukou	37.37	42.26	41.95
Age	26.95	24.08	24.39
Income	29.27	31.42	31.25
Region	2.5	1.18	1.2

Note: (1) Values in Table 12 represent relative influence among independent variables. (2) Contribution of region includes the east and central component. (3) Measured by percent.

rare in rural areas. Furthermore, the hukou system, which registers urban and rural households separately, leads people to seek an education where one's residence is registered. Consequently, rural students are disadvantaged regarding educational attainment, especially with regards to good quality education which benefits the achievement of a higher degree. In summary, the hukou status is deeply rooted in an individual's educational attainment, which highlights the role China's rural–urban divide in shaping people's labor-market return (Fu and Ren, 2010).

Next, income factor is a secondary contribution to education inequality, which means that social stratification division plays an important role in inequality. At this point in time, households with social stratification advantages translate their social power and economic resources into better chances of education for their children. This kind of mechanism has been widely acknowledged by many studies, such as Treiman and Yip (1989).

In practice, there are always two patterns which represent the unequal effect of social stratification. The first is dominant exclusion, which means advantaged groups can send their children to a good quality school. According to our investigation, many key schools always have two admission procedures in China, the elite one is for students with high scholastic achievement and no extra fees need to be paid except for tuition fees, while the inferior one is for students with low scholastic achievement where some extra expenses for selecting school must be paid. So the power from advantaged groups has been translated into educational privileges. Another pattern is recessive exclusion, which means disadvantaged groups have a lower ability for risk taking or low estimation of educational return when deciding whether to enter a higher level school (Breen and Goldthorpe, 1997). Nowadays in China, with the situation of deep income inequality and some employment discrimination based on family background, many students from disadvantaged groups abandon opportunities for higher educational achievement when considering opportunity cost. This kind of pattern seems not to build a barrier for disadvantaged groups, but forces disadvantaged groups to draw a 'rational choice' under huge pressures from economic or institutional factors.

Further, an age factor contributes to education inequality. The reason may lie in that older people usually have less education than the young. Education expansion policies have improved the young's educational attainment in last decades, especially the compulsory education policy of nine years which ensures most children, whether urban or rural, could receive a basic education.

In addition, gender difference does not represent a major factor in overall education inequality. Equal educational opportunities for females were widely available so that each female student could get a basic education. Although there is still an educational gap between males and females, that partly results from the educational system and other causes such as cultures, customs, and family notions.

At last, influences from regional differences are also not very big. Each province has its own resource endowment so that various educational policies should be designed and applied. However, that does not mean regional differences no longer exist. Because of unbalanced economic development between different areas of the country, provinces in the east are able to invest more funds toward education, and such a promotional effect will be significant after an initial lag phase. In addition, the disparity in economic development always causes more talent to flow into the developed regions. This is not only a main reason of educational gaps between regions, but also a negative incentive for the developing areas to invest in education.

5. Conclusion

The purpose of this study is to analyze the situation and formation mechanism of educational inequality in China. At the beginning, we adopted the education Gini coefficient to measure the education inequality of China. The results show that the education expansion policy in last decades has contributed to a sharp decrease in education inequality and an improvement of average years of schooling (AYS). And there exists a universal law which states that the higher AYS level achieved, the lower the education inequality will be. Overall, educational development gaps among regions are still deep.

Next, based on China General Social Survey (CGSS) data in 2006 and Gini coefficient decomposition method by subgroups, we decomposed education inequality for the variables of urban–rural division, gender difference, regional disparity, social stratification division, and age. The findings are that the educational development gap between urban and rural is so deep because people from urban areas have more opportunities for higher educational achievement. When decomposing as gender difference, although educational disparity between male and female has been sharply decreased, gaps still exist and deserve attention, especially in areas of rural poverty. Results based on decomposing as regional disparity indicate the central government should increase educational transfer payments to less developed provinces, and establish a financial system governed by a provincial institution. Furthermore, it is worth mentioning that a between-group component contributes to total education inequality the most when including influences from the social stratification division. More specifically, stratum differentiation deepens educational gaps between groups with the background of income inequality. At

last, decomposition results of age indicate that increasing educational attainment for the young plays a key role in reducing education inequality.

In order to know which of the factors above contribute to education inequality most, Shapley value decomposition based on regression is used. Results show that urban–rural division contributes to education inequality most, followed in decreasing significance by social stratification division, age, gender disparity and regional gaps. It is worth noting that the hukou system has formed the disparity in educational resources and investment between urban and rural areas. Moreover, with the background of increasing income inequality, social stratification division would strengthen such an already negative effect. With the combination of system discrimination and power, a relatively dominant class is enjoying greater educational opportunities and good quality educational resources, but the disadvantaged are excluded. This is a rising trend of education inequality in China. Unfortunately, the Chinese government has paid minimal attention to such education inequality; some possible corrective measures have not been taken.

Now China is experiencing rapid economic development, as the government is trying to eliminate the urban–rural division during the process of urbanization and industrialization. Predictably, more and more peasants and their offspring will swarm into cities in the decades ahead. However, insufficient investment in education and unequal distribution of good quality education might be a primary reason for the migration, further causing education inequality in the long term. But more than that, increasing income inequality and deepening stratum differentiation will make education reform challenging. Designing a scientific and reasonable mechanism should focus on the disadvantaged group's educational desire, and emphasize balanced development in education in all areas.

Acknowledgements

This study is supported by Chinese National Natural Science Foundation (No. 70703038), Chinese National Education Science Project (No. EFA080248), and Key Projects of Philosophy and Social Sciences Research, Ministry of Education (No. 13JZD023).

References

Barro, R.J., Lee, J.W., 1993. International comparisons of educational attainment. Journal of Monetary Economics 32 (3), 363–394.
Blinder, A., 1973. Wage discrimination: reduced form and structural estimates. Journal of Human Resources 8 (3), 436–455.
Breen, R., Goldthorpe, J.H., 1997. Explaining educational differentials: towards a formal rational action theory. Rationality and Society 9 (3), 275–305.
Breusch, T.S., Pagan, A.R., 1979. A simple test for heteroscedasticity and random coefficient variation. Econometrica 47 (5), 1287–1294.
Dong, Q., Li, X., Yang, H., Zhang, K., 2008. Gender inequality in rural education and poverty. Chinese Sociology and Anthropology 48 (4), 64–78.
Ebert, U., 1988. Measurement of inequality: an attempt at unification and generalization. Social Choice and Welfare 5 (2/3), 147–149.
Fu, Q., Ren, Q., 2010. Educational inequality under China's rural–urban divide: the hukou system and return to education. Environment and Planning 42 (3), 592–610.
Hannum, E., 1999. Political change and the urban–rural gap in basic education in China, 1949–2009. Comparative Education Review 43 (2), 193–211.
Hannum, E., Wang, M., 2006. Geography and educational inequality in China. China Economic Review 17, 253–265.
Lambert, P.J., Aronson, J.R., 1993. Inequality decomposition analysis and the Gini coefficient revisited. Economic Journal 103, 1221–1227.
Lin, C-H.A., Yang, C.-H., 2009. An analysis of educational inequality in Taiwan after the higher education expansion. Social Indicators Research 90 (2), 295–305.
Mookherjee, D., Shorrocks, A., 1982. A decomposition analysis of the trend in UK income inequality. Economic Journal 92 (368), 886–902.
Oaxaca, R., 1993. Male-female wage differences in urban labour markets. International Economic Review 14 (3), 693–709.
Pyatt, G., 1976. On the interpretation and disaggregating of Gini coefficients. Economic Journal 86 (342), 243–255.
Qian, X., Smyth, R., 2005. Measuring regional inequality of education in China: widening cost-inland gap or widening rural–urban gap? ABERU Discussion Paper, No. 12. Department of Economics, Monash University, Australia.
Ram, R., 1990. Educational expansion and schooling inequality: international evidence and some implications. Review of Economics and Statistics 72 (2), 266–274.
Soltow, L., 1960. The distribution of income related to changes in the distribution of education, age and occupation. Review of Economics and Statistics 42 (4), 450–453.
Silber, J., 1989. Factor components, population subgroups and the computation of the Gini index of inequality. Review of Economics and Statistics 71 (1), 107–115.
Shorrocks, A., Wan, G., 2005. Spatial decomposition of inequality. Journal of Economic Geography 5 (1), 59–81.
Treiman, D.J., Yip, K., 1989. Educational and occupational attainment in 21 countries. In: Kohn, M.L. (Ed.), Cross-national Research in Sociology. Sage, Beverly Hills, CA, pp. 373–394.
Tomas, V., Wang, Y., Fan, X., 2003. Measuring education inequality: Gini coefficients of education for 140 countries, 1960–2000. Journal of Education Planning and Administration 17 (1), 5–33.
Tsang, M.C., 1996. Financial reform of basic education in China. Economics of Education Review 15 (4), 423–444.
Wan, G., 2002. Regression-based inequality decomposition: pitfalls and solution procedure. WIDER Discussion Paper.
Wan, G., 2004. Accounting for income inequality in rural China. Journal of Comparative Economics 32 (2), 348–363.
White, H., 1980. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. Econometrica 48 (4), 817–838.
World Bank, 2010. World Development Indicators. World Bank, Washington, DC.
Yitzhaki, S., Lerman, S., 1991. Income stratification and income inequality. Review of Income and Wealth 37 (3), 313–329.
Yang, J., Huang, X., Li, X., 2009. Education inequality and income inequality: an empirical study on China. Frontier of Education in China 4 (3), 413–434.
Zhang, J., Li, T., 2002. International inequality and convergence in education attainment, 1960–1990. Review of Development Economics 6 (3), 383–392.