大学院研究年報　第52号　2023年２月

An Analysis of Regional Educational Disparities in China

Li Feng *

Abstract

This study analyzes regional educational disparities in China by measuring the Gini coefficient of the population's Average Years of Schooling (AYS) and Percentage of Graduates of junior secondary schools entering senior secondary schools (PG) in coastal and inland provinces, in rural and urban areas and in China as a whole. I calculate the Gini coefficient for AYS using 2010 and 2018 micro-sample data from the China Family Panel Studies (CFPS) and for PG using 2010 and 2018 aggregate data from the 2011 and 2019 China Education Yearbooks. This paper adopts the decomposition method of the Gini coefficient to examine educational inequality within China's coastal and inland provinces (and in its rural and urban areas) as well as between the coastal and inland provinces and rural-urban areas. It compares the results of this research with those of previous research such as Qian and Smyth (2008), and observes the changes in regional educational inequality in China over the course of 28 years. Then, it uses a regression analysis to elaborate on regional educational disparities in China. Finally, it takes a Shapley decomposition approach based on the results of regression analysis to examine contributing factors to disparity and the extent to which they contribute to educational disparity as a whole in China.

Contents

＊　リ　ホウ　　経済学研究科経済学専攻博士課程後期課程

Ⅰ Introduction

Economists tend to be big fans of education, which is perhaps not surprising given how much of it they consume and how well their textbooks can do. Alfred Marshall, writing in 1873, hoped that education would help erase the "distinction between working men and gentlemen". Gary Becker of the University of Chicago reimagined education as an investment in "human capital" that would earn a return in the market much like other assets. Harvard University's Greg Mankiw, whose books have educated more than most, once calculated that differences in human capital between countries could account for much of their otherwise inexplicable differences in prosperity. Education is not only beneficial to individuals, but to the countries they call home. It is the fundamental to the growth of economics (Lucas, 1988). It significantly reduces the probability of incarceration and arrest and raises the opportunity costs of crime and of time spent in prison. Education can also make individuals less impatient or more risk averse, further reducing the propensity to commit crimes (Lochner and Moretti, 2004).

Education plays a central role in the modern labor market. Hundreds of studies in many different countries and time periods have confirmed that better-educated individuals earn higher wages, experience less unemployment, and work in more prestigious occupations than their less-educated counterparts (Card, 1999; Psacharopoulos, 1985, 1994). This is directly associated with returns to education. Students who are compelled to attend school longer by compulsory schooling laws earn higher wages as a result of their extra schooling (Angrist and Keueger, 1991). China's education returns for additional years of schooling generally increased from 2.6% in 1989 to 7.9% in 2009 (Fleisher and Wang, 2004; Fu and Ren, 2010; Li and Zhang, 1998). This is one of the reasons why Chinese parents invest so heavily in their children's education.

The *hukou* (household register)[1] system, instituted in cites in 1951 and extended to rural areas in 1955, is critical to understandings of education of China. In the early years of the system, it served largely as a monitoring, rather than control, mechanism for population migration and movements. In fact, the early 1950s was a period of relatively free movement into and out of the cities and throughout the countryside. The constitution promulgated in 1954 even guaranteed citizens' rights of free residential choice and migration. However, as influxes of peasants into cities escalated and began to be a serious burden on urban infrastructure, the central government tried various measures to stop what it called "blind flows" of rural labor. This culminated in the promulgation of China's first set of *hukou* legislations by the National People's Congress in 1958 (Chan and Zhang, 1999; Wu and Treiman, 2004; Wang, 2004). Next, China's *hukou* system was formally implemented by the government, assigning every Chinese citizen an agricultural *hukou* type or a nonagricultural *hukou* type. In addition, Chinese citizens were assigned a *hukou* location, that is, where they are registered or where their *hukou* is. The *hukou* location was normally one's place of birth (Chen and Fan, 2016).

Why is *hukou* significant? *Hukou* registration not only provided the principal basis for establishing identity, citizenship, and proof of official status, but it was essential in every aspect of daily life: "Without registration, one cannot establish eligibility for food, clothing or shelter, obtain employment, go to school, marry or enlist in the army" (Cheng and Selden, 1994). In terms of the labor market and employment, studies suggest that migrating rural holders suffer from *hukou*-based labor market discrimination (Jiang, Lu and Sato, 2012). Arrow (1973) defines labor market discrimination as the valuation in the marketplace of a worker's personal characteristics unrelated to worker productivity. The average hourly wage for local urban *hukou* workers is 7.8 yuan, while it is only 4.4 yuan for rural migrants (Lee, 2012).

In terms of education, almost half of the rural migrants willing to convert their rural *hukou* to urban *hukou* cited their children's education (about 49 percent) as the main reason, followed by urban social benefits (about 15 percent), urban living environment (about 14 percent), and employment opportunities (about 14 percent) (Chen and Fan, 2016). Education for one's children is one of the primary reasons a person might seek an urban *hukou*. One of the many problems facing migrants in cities is that even if their child is born in the city, that child will not be able to receive the same quality of education if they are registered elsewhere. Currently in cities where large numbers of young migrants are registered with rural *hukous,* such as Shanghai, those children are only allowed to go to schools for grades 1–9. Even if these students do well in school, there is a very slim chance that they would be accepted into public schools or even universities. If a child is born with a Shanghai urban *hukou*, not only are they more likely to go to a good high school, but their entrance exam scores do not need to be as high for entrance into some of the larger Chinese universities such as Peking University and Shanghai Jiao Tong University (Kryshak, 2016; Ling, 2015).

Ⅱ Previous Studies

A number of studies investigate educational disparities in China. Some of these studies aim to measure and understand the rural-urban gap in cognitive ability in China: on average, urban students score approximately 1.4 points (about 17 percent) higher on cognitive ability tests than rural students, which is equivalent to about 37 percent and about 41 percent of the standard deviations for urban and rural students, respectively (Zhao et al., 2017). This gap in cognitive ability is paired with gaps in educational access as well. Disparities in access to education between rural and urban areas rather than between coastal and inland provinces represent the primary root of educational inequality in China (Qian and Smyth, 2008). As Wu (2011) suggests, rural people in China are highly disadvantaged in years of schooling attained, and this disparity has been consistent over time.

Children from poorer backgrounds do less well in a number of dimensions that their peers (Gregg and Machin, 2000). In terms of completed education, children from low-income households go on to leave full-time education much earlier, and with fewer formal qualifications, than their more affluent counterparts (Blanden and Gregg, 2004). Thus, household background and income are crucial for children. In China, children from households where the parents are more highly educated and work in occupations involving either expertise or administrative authority develop higher educational aspirations than do children from less well-educated and blue-collar or lower-white-collar backgrounds (Broaded and Liu, 1996). In other words, family economic conditions influence children's and adolescents' education through home educational resources and parental involvement (Cai and Wu, 2019).

Many studies have also shown that gender plays a crucial role in educational attainment. For example, there is strong evidence of a gender selection in which poorly performing girls drop out in primary school while poorly performing boys do not begin to drop out in earnest until junior secondary school (Brown and Park, 2002). The reality of dropping out of school has also been studied by other researchers who have found high rates of school attendance and retention among boys but high rates of withdrawal from school among girls in poor regions. Among adults, the rate of illiteracy is much higher in women than in men, and the average number of years of education is much greater for men than women (Dong et al., 2008). Thanks to, the compulsory education system and gender equity promotions, the gender gap in educational attainment has been greatly reduced in recent decades, but 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 pursuing education. In short, though educational disparities between men and women have decreased sharply, differences still remain and deserve further attention (Yang et al., 2014).

Age is usually organized as an important factor when considering educational disparities in China. Older people have comparatively lower rates of secondary and university education. As part of the Cultural Revolution (1966-76), universities were shut down altogether in 1966, and exclusively political criteria for admission were employed when they reopened in 1972; similar interventions occurred at the secondary school level as well. The massive state intervention of the Cultural Revolution drastically reduced the educational attainment of a whole cohort of the population and attenuated any educational advantage of coming from an educated family. When children of Cultural Revolution were ages 25-35, thereby not allowing for the possibility of catch-up at later points in time (Deng and Treiman, 1997). However, this is not the main cause for age disparities in education in China. Older people have comparatively lower school rates of secondary and university education; only 5% of eligible citizens where enrolled in university in 1997, and only 30% of eligible citizens were enrolled in secondary school in 1984. In other words, China had lower school enrollment in the past. With the Reform and Opening up in 1978 and the introduction of Compulsory Education in 1986, people born in the 1990s had higher school enrollment. Age, therefore, is one of the factors considered in this study.

The final facto considered in this study is regional division. For example, the link between geography and access to primary school has greatly increased across cohorts, as the few without access to primary school are ever more concentrated in poor areas (Hannum and Wang, 2006). Moreover, between regions, there are stark differences not only in educational attainment, but also in the allocation of educational resources. For instance, because of imbalances in economic development across different areas of the country, provinces in the east are able to invest more toward education, the effects of which are significant after an initial lag phase (Yang et al., 2014). Indeed, Peng et al. (2020) suggest that provincial-level variations in the child population and child dependency ratios have made access to educational resources even more unequal, given unequal financial capacities at the provincial level. Poorer provinces have higher child dependency ratios and less economic development; jointly, these two factors lead to limited educational resources (Peng et al., 2020). Thus, region is also considered crucial for studies of educational disparities in China.

This study mainly refers to the methods of Qian and Smyth (2008) and Yang et al., 2014. Using the 2011 and 2019 China Education Yearbooks, it compares the results of this research with those of the previous studies (Qian and Smyth, 2008). In order to find out the contributing factors to educational disparities in China and their effects, the study uses the Shapley decomposition approach based on regression analysis from Yang et al. (2014).

Ⅲ The Educational Context in China

After 40 years of social reform and an opening up policy initiated by Deng Xiaoping, China has made remarkable achievements in education. Figure 1, for instance, shows decreases in the illiterate population from 2000 to 2018. To examine the educational context in China more closely, this study further focuses on *hukou* (urban and rural), income (in different geographical areas), gender and the expenditure of education in geographical areas (east, central, west).

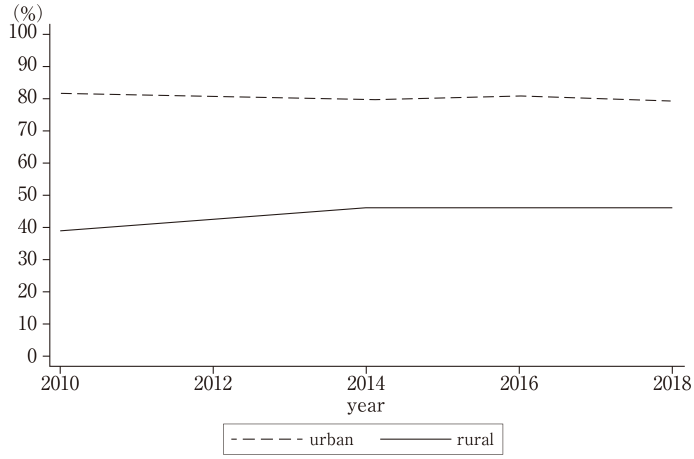
Chapters 1 and 2 consider continued problems with *hukou* (urban and rural) system in modern China, especially disparities in education between urban and rural areas: the PG index in rural areas was almost 43% from 2010 to 2018, while in urban areas it was almost double, 80% (Figure 2).

Income affects individual human capital investment levels directly, but it also affects them indirectly, by reinforcing disparities in education. Income inequality leads to educational inequality, while the reduction of educational equality does not lead to decreases in income inequality; there is no simple casual effect between the two. However, education expansion is beneficial for reducing both educational and income inequality (Yang, Huang and Li, 2009). Figure 3 shows the drastic per capita disposable income disparities between urban and rural areas in China. Per capita disposable income is much higher in urban areas than in rural areas.

Figure 1 The rate of the illiterate population in China, ages 15 and older, from 2000 to 2018

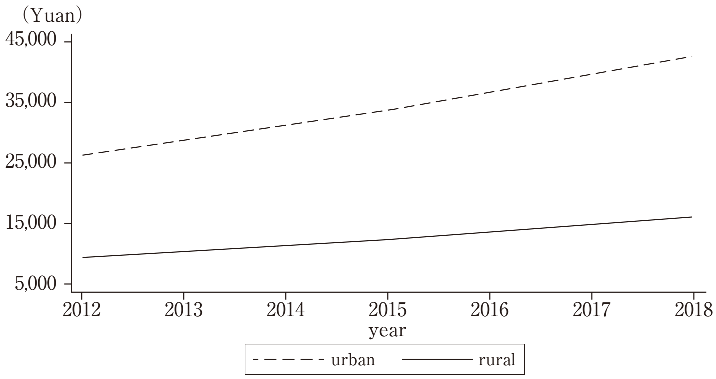
Source: National Bureau of Statistics, China Education Yearbook (1999-2019).

Figure 2 PG index (2010-2018)



Source: National Bureau of Statistics, China Education Yearbook (2011-2018).

Figure 3 Per capita disposable income in China

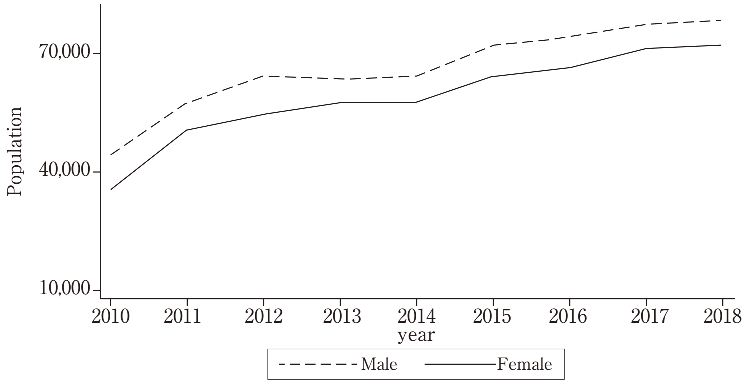


Source: National Bureau of Statistics, China Education Yearbook (2013-2018).

Figure 4 highlights gendered disparities in higher education in modern Chinses society. Higher education attainment is higher among males than among their female counterparts (Figure 4).

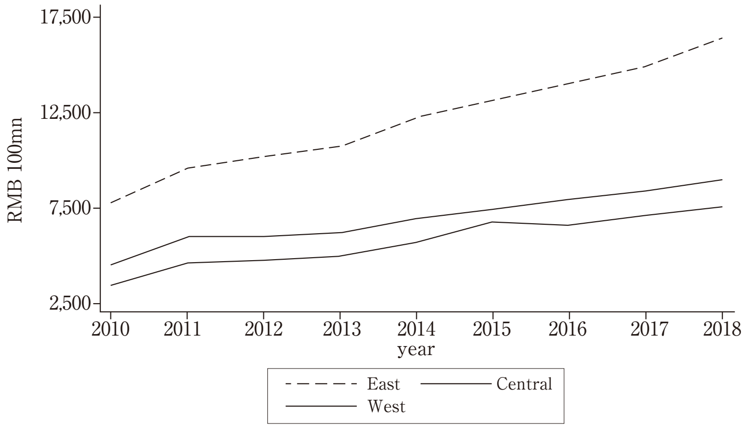
Western and Central China tend to greater educational inequality than Eastern China. This phenomenon follows the pattern of uneven economic development among provinces in China. Therefore, the government should pay more attention to the west and central areas when considering educational investment and economic policy. Educational resources and overall resource distribution are both particularly scarce in the west (Yang, Huang and Li, 2009). This study uses the data on education expenditure on the part of the NBS to compare the education expenditure in different areas of China. Education expenditures in the east are much higher than in the central and western provinces (Figure 5).

Figure 4 Population having attained higher education in China



Source: National Bureau of Statistics, China Education Yearbook (2011-2019).

Figure 5 Education expenditures by provincial governments in China



Source: National Bureau of Statistics, China Education Yearbook (2011-2019).

Ⅳ Data and Analytical Methods

Two data sources are used in this study. First I use the China Education Yearbook of the National Census in 2010 and 2018 to calculate AYS (the average years of schooling in a population) by province and PG (the percentage of graduates of junior secondary schools entering senior secondary schools) by rural-urban area (NBS, 2011, 2019).

Second, to calculate AYS and perform the regression analysis, I use the China Family Panel Studies (CFPS), a nationally representative, biannual longitudinal survey of Chinese communities, families, and individuals produced by the Institute of Social Science Survey (ISSS) of Peking University, China. The CFPS collects, family, and community-level longitudinal data covering a wide range of economic activities, education outcomes, family dynamics and relationships, migration, and health. In the 2010 baseline survey, the CFPS successfully interviewed 14,960 households and 42,590 individuals, with an approximate response rate of 79%. Respondents are tracked through annual follow-up surveys.

Analytical Methods

In this study, I use two approaches to educational attainment: the average years of schooling of the population (AYS) of a given province and the percentage of graduates of junior secondary schools entering senior secondary schools (PG) for each province separately in rural and urban areas. The formula to calculate AYS is:

$$\mathrm{AYS} = \frac{5\mathrm{H}_1+8\mathrm{H}_2+10.5_3+14\mathrm{H}_4}{\mathrm{POP}} \tag{1}$$

Here H_t is the number of people for whom t is the highest level of schooling attained; j=1, for primary, 2 for junior secondary school, 3 for senior secondary school and secondary technical school and 4 for college and above. The duration of the jth level of schooling is adopted from Wang and Yao (2003). POP is the total population. The formula used to calculate PG is:

$$\mathrm{PG}\ (\%) = \frac{\text{New student enrolment of secondary schools}}{\text{Graduates of junior secondary schools}} \times 100\% \tag{2}$$

To measure the relative inequality of the schooling attained, I develop an indicator for the educational Gini coefficient, adopting an indirect method originated by Tomas et al. (2001) based on educational attainment data. The education Gini formula is:

$$G^E = \frac{1}{\mu N(\mathrm{N}-1)} \sum_{i>j} \sum_{j} | x_i - x_j | \tag{3}$$

Here G^E is the education Gini index; μ is the mean value of AYS or PG of the total sample; N is the total number of observations in the sample; x_i and x_j are AYS (of a given province) or PG (of urban or rural areas of a province).

In decomposing the Gini coefficient I draw from Zhang and Li (2002):

$$\mathrm{G} = p_1^2\left(\frac{\mu_1}{\mu}\right)G_1 + P_2^2\left(\frac{\mu_2}{\mu}\right)G_2 + G_B \tag{4}$$

Here, P_i, represents the proportion of the population of subgroup i; μ_i, is mean AYS or PG of group i, μ is mean AYS or PG for the whole country; and G_i is the Gini coefficients for subgroup i (i = 1, 2). G_B is the between-group contribution to total inequality.

An alternative measure of educational inequality is a polarization index, as defined by Zhang and Kanbur (2001):

$$\mathrm{P} = \frac{\text{Between–group inequality}}{\text{Within–group inequality}} \tag{5}$$

Rather than decompose the Gini index to reveal the contributions of different groupings to overall educational disparity, this strategy the magnitude of polarization by the ratio of inequality between and within each grouping.

Shapley Decomposition (Shorrocks, 1999)

Consider a statistical indicator I whose value is completely determined by a set of m contributory factors, X_k, indexed by $k = K = (1, 2, \dots, m)$ so that we may write

$$I = f\,(X_1, X_2, \dots, X_m) \tag{6}$$

where $f\,(\cdot)$ describes the underlying model. In the applications examined later, the indicator I will represent the overall level of poverty or inequality in the population, or the change in poverty over time. The factor X_k may refer to a conventional scalar or vector variable, but other interpretations are possible and often desirable; for the moment it is best regarded as a loose descriptive label capturing influences like "uncertain return to investments", "difference in household composition" or "supply-side effects".

In what follows, we imagine scenarios in which some or all of the factors are eliminated, and use F (S) to signify the value that I takes when the factors X_k and $k \notin S$ have been dropped. As each of the factors is either present or absent, it is convenient to characterize the model structure (K, F) in term of the set of factors (or , more accurately, "factor indices"), K, and the function $F{:}\{S|S\subseteq K\}\rightarrow\mathbb{R}$. Since the set of factors completely accounts for I, it will also be convenient to assume throughout that $F(\phi)=0$: in other words, that I is zero when all the factors are removed.

A decomposition of (K, F) is a set of real values C_k, $k\in K$, indicating the contribution of each of the factors. A decomposition rule C is a function that yields a set of factor contributions

$$C_k = C_k\,(K, F),\ k\in K \tag{7}$$

For any possible model (K, F).

In seeking to construct a decomposition rule, several considerations come to mind. First, it is should be symmetric (or anonymous) in the sense that the contribution assigned to any given factor should not depend on the way in which the factors are labelled or listed. Secondly, that decomposition should be exact (and additive), so that

$$\sum_{k\in K} C_k\,(k, F) = F\,(k) \tag{8}$$

for all (K, F). When condition (8) is satisfied, it is meaningful to speak of the proportion of observed inequality or poverty attributable to factor k.

It is also desirable that the contributions of the factors can be interpreted in an intuitively appealing way. In this respect, the most natural candidate is the rule that equates the contribution of each factor to its (first round) marginal impact

$$M_k\,(K, F) = F\,(K) - F\,(K \setminus \{k\}),\ k \in K \tag{9}$$

This decomposition rule is symmetric, but will not normally yield an exact decomposition. A second possibility is to consider the marginal impact of each the factors when they are eliminated in sequence. Let $\sigma = (\sigma_1, \sigma_2, \dots \sigma_m)$ indicate the order in which the factors are removed, and let $S(\sigma_r, \sigma)=\{\sigma_i \mid i > r\}$ be the set of factors that remain after factor σ_r has been eliminated. Then the marginal impacts are given by

$$C_K^\sigma = F(S(k, \sigma) \cup \{k\}) - F(S(k, \sigma) = \Delta_k F(S(k, \sigma), k \in K \quad (10)$$

where

$$\Delta_k F(S) \equiv F(S \cup \{k\}) - F(S), S \subseteq K \setminus \{k\} \quad (11)$$

is the marginal effect of adding factor k to the set S. Using the fact that $S(\sigma_r, \sigma) = S(\sigma_{r+1}, \sigma) \cup \{\sigma_{r+1}\}$ for r=1, 2, …$m-1$, we deduce

$$\begin{aligned} \textstyle\sum_{k \in K} C_k^\sigma = \sum_{r=1}^m C_{\sigma_r}^\sigma = \sum_{r=1}^m [F(S(\sigma_r, \sigma) \cup \{\sigma_r\}) - F(S(\sigma_r, \sigma))] \\ F(S(\sigma_1, \sigma) \cup \{\sigma_1\}) - F(S(\sigma_m, \sigma)) = F(K) - F(\phi) = F(K) \end{aligned} \quad (12)$$

The decomposition (10) is therefore exact. However, the value of the contribution assigned to any given factor depends on the order in which the factors appear in the elimination sequence σ, so the factors are not treated symmetrically. This "path dependence" problem may be remedied by considering the m! possible elimination sequences, denoted here by the set Σ, and by computing the expected value of C_k^σ when the sequences in Σ are chosen at random. This yields the decomposition rule C^s given by

$$\begin{aligned} C_K^S(K, F) &= \frac{1}{m!}\sum_{\sigma \in \Sigma} C_k^\sigma = \frac{1}{m!}\sum_{\sigma = \Sigma} \Delta_K F(S(k, \sigma)) \\ &= \sum_{s=0}^{m-1} \sum_{\substack{S \subseteq K \setminus \{k\} \\ |S|=s}} \frac{1}{m!} \sum_{\substack{\sigma \in \Sigma \\ S(k, \sigma)=S}} \Delta_k F(S) \\ &= \sum_{s=0}^{m-1} \sum_{\substack{S \subseteq K \setminus \{k\} \\ |S|=s}} \frac{(m-1-s)!s!}{m!} \Delta_k F(S) \end{aligned} \quad (13)$$

Using $\pi(s, m-1) = (m-1-s)!s!/m!$ to indicate the relevant probability, Eq. 13 is expressed more succinctly as

$$C_k^s(K, F) = \sum_{S \subseteq K \setminus \{k\}} \pi(S|, K \setminus \{k\}) \Delta_k F(S) = \varepsilon_{S \subseteq K \setminus \{k\}} \Delta_k F(S), k \in K \quad (14)$$

where $\varepsilon_{S \subseteq L}$ is the expectation taken with respect to the subsets of L.

From Eq.12 it is clear that C^S is an exact decomposition rule, and also one that treats the factors symmetrically. Furthermore, the contribution can be interpreted as the expected marginal impact of each factor when the expectation is taken over all the possible elimination paths.

Expression 13 will be familiar to many readers, since it corresponds to the Shapley value for the cooperative game in which "output" or "surplus" $F(K)$ is shared amongst the set of "input" or "agents" K. The application to distributional analysis is quite different from the context in which the Shapley value was conceived, and the results therefore need to be reinterpreted. Nevertheless, it seems convenient and appropriate to refer to Eq.13 as the Shapley decomposition rule.

V The Results of the Quantitative Analysis

Educational Attainment Measured by Average Years of Schooling (NBS)

We have already calculated the Gini coefficient to analyze the Average Years of Schooling of the population (AYS) and the Percentage of Graduates of junior secondary schools entering senior secondary schools (PG) in 2010 and 2018, using the 2010 and 2018 aggregate data from the China Education Yearbook. The results described in Table 1 indicate the changes from 2010 to 2018.

Table 1 shows that the Gini coefficient in China as a whole increased by 10.4% from 0.0663 in 2010 to 0.0732 in 2018 over the 9- year study period. The Gini coefficient in the inland provinces decreased by 2.9% from 0.0428 to 0.0416, and the Gini coefficient in the coastal provinces increased by 20.2%.[2)]

The results of decomposition of the Gini coefficient described in Table 2 show that from 2010–2018 the contribution of inequality in the coastal provinces to that in China as a whole increased by 55.7%, from 20.4% to 31.8%, while the contribution of inequality within the inland provinces decreased by 31.5%, from 24.5% to 16.8%. The contribution of inequality within coastal-inland provinces decreased by 6.6%, from 55.1% to 51.5%.

The polarization index of educational disparities confirms the results from the decomposition of Gini

Table 1 AYS and Gini Coefficients in 2010 and 2018 (NBS Data)

	All Provinces			Coastal Provinces			Inland Provinces		
	2010	2018	Δ(%)	2010	2018	Δ(%)	2010	2018	Δ(%)
AYS (years)	8.1484	8.8245	8.30	8.5277	9.2388	8.34	7.9288	8.5847	8.27
Gini	0.0663	0.0732	10.41	0.0917	0.1102	20.21	0.0428	0.0416	-2.86

Sources: Calculated based on NBS, China Education Yearbook (2011 and 2019).

Table 2 Inequality Decomposition by Coastal and Inland Provinces (AYS/NBS data)

(unit: %)

	Within Coastal Provinces	Within Inland Provinces	Coastal-Inland Gap
2010	20.4	24.5	55.1
2018	31.8	16.8	51.5
Δ(%)	55.7	-31.5	-6.6

Sources: Calculated based on NBS, China Education Yearbook (2011 and 2019).

Table 3 Polarization Index of Educational Disparities between Coastal and Inland Provinces (AYS/NBS Data)

	Between-Group Inequality	Within-Group Inequality	B/W
2010	0.0365	0.0298	1.2276
2018	0.0377	0.0355	1.0601

Sources: Calculated based on NBS, China Education Yearbook (2011 and 2019).

coefficients described in Table 3. It shows that the ratio of between-groups to within-groups disparities decreased from 1.2 to 1.1.

Educational Attainment Measured by Average Years of Schooling (CFPS)

In order to find out whether 2010 and 2018 aggregate data from the China Education Yearbook and 2010 and 2018 micro-data from China Family Panel Studies (CFPS) share the same tendency of results, this study uses the same methods described in chapter 4 to calculate the results of CFPS, described in Table 4.

Table 4 shows that the Gini coefficient in China as a whole increased by 12.2%, from 0.0949 in 2010 to 0.1065 in 2018, over the 9-year period. The Gini coefficient in the coastal provinces increased by 1.9%, from 0.1241 to 0.1265. Meanwhile, the Gini coefficient in the inland provinces increased by 9.6%.

The results of decomposition of the Gini coefficient described in Table 5 show that the contribution of inequality in the coastal provinces to that of China as a whole decreased by 23.1% from 31.9% to 24.6% during the period of 2010–2018, while the contribution of inequality in the inland provinces to China overall increased by 13.3%, from 19.4% to 22%. The contribution of the coastal-inland inequality increased by 9.9%, from 48.7% to 53.5%.

The polarization index of educational disparities confirms the results from the decomposition of Gini coefficients described in Table 6. It shows that the ratio of between-group to within-group disparities

Table 4 AYS and Gini Coefficients in 2010 and 2018 (CFPS Data)

	All Provinces			Coastal Provinces			Inland Provinces		
	2010	2018	Δ(%)	2010	2018	Δ(%)	2010	2018	Δ(%)
AYS (years)	8.1458	9.1115	11.86	8.3525	9.4865	13.58	7.9833	8.8169	10.44
Gini	0.0949	0.1065	12.22	0.1241	0.1265	1.93	0.0716	0.0785	9.64

Sources: Calculated based on China Family Panel Studies (2010 and 2018).

Table 5 Inequality Decomposition by Coastal and Inland Provinces (AYS/CFPS Data)

(unit: %)

	Within Coastal Provinces	Within Inland Provinces	Coastal-Inland Gap
2010	31.9	19.4	48.7
2018	24.6	21.9	53.5
Δ(%)	-23.1	13.3	9.8

Sources: Calculated based on China Family Panel Studies (2010 and 2018).

Table 6 Polarization Index of Educational Disparities between Coastal and Inland Provinces (AYS/CFPS Data)

	Between-Group Inequality	Within-Group Inequality	B/W
2010	0.0523	0.0426	1.2276
2018	0.0548	0.0517	1.0599

Sources: Calculated based on China Family Panel Studies (2010 and 2018).

decreased from 1.2 to 1.1.

Educational Attainment Measured by PG Approach (NBS)

Based on the results of PG analysis described in Table 7, the Gini coefficient for China as a whole decreased by 38.2%, from 0.1572 in 2010 to 0.0943 in 2018. The coefficient in the coastal provinces decreased by 21.6%, from 0.1354 to 0.1061. The Gini coefficient in the inland provinces also decreased by 43%, from 0.1591 to 0.0907. For the same period, while the Gini coefficient in the urban areas increased by 4.3%, from 0.1440 to 0.1502, the Gini coefficient in the rural areas decreased by 17.8%, from 0.1988 to 0.1635.

Looking at the results of the decomposition analysis of the Gini coefficient described in Table 8, the contribution of rural and urban areas to overall educational inequality indicated a dramatic decline of 63.6% from 39.2% to 14.3%. However, the contributions of inequality within rural and urban areas to overall inequality increased by 14.7%, from 31.1% to 35.7%, and by 68.6%, from 29.7% to 50%,

Table 7 PG and Gini Coefficients in 2010 and 2018 (NBS Data)

	All Provinces		Rural Areas within Provinces		Urban Areas within Provinces		Coastal Provinces		Inland Provinces	
	2010	2018	2010	2018	2010	2018	2010	2018	2010	2018
PG	51.16	59.83	38.86	46.55	83.48	79.71	53.86	60.00	49.59	59.74
Gini	0.1527	0.0943	0.1988	0.1635	0.1440	0.1502	0.1354	0.1061	0.1591	0.0907

Sources: Calculated based on NBS, China Education Yearbook (2011 and 2019).

Table 8 Inequality Decomposition by Rural and Urban Areas and by Coastal and Inland Provinces (PG/NBS Data)

(unit: %)

	Contribution of the Rural/Urban gap to Overall Education Inequality			Contribution of the Coastal/Inland Gap to Overall Educational Inequality		
	Within Rural Areas	Within Urban Areas	Rural-Urban Gap	Within Coastal Areas	Within Inland Areas	Coastal/Inland Gap
2010	31.1	29.7	39.2	14.9	36.4	48.6
2018	35.7	50	14.3	18.3	34.3	47.4
Δ(%)	14.7	68.6	-63.6	22.8	-5.9	-2.6

Sources: Calculated based on NBS, China Education Yearbook (2011 and 2019).

respectively. The contributions of inequality within coastal and inland areas to overall inequality decreased slightly by 2.6%, from 48.6% to 47.4%: within coastal provinces, the contribution of inequality to overall inequality increased by 22.8%, from 14.9% to 18.3%, while for inland provinces it declined by 5.9%, from 36.4 to 34.3.

To find out more about the disparities between rural and urban areas of the coastal and inland provinces, this study further decomposes the Gini coefficient within each group. The results are described in Tables 9 and 10. Table 9 shows the decomposition of rural-urban educational disparities for coastal and inland provinces. The results show that whether educational disparities in coastal for rural-urban or educational disparities of inland for rural-urban decreased from 2010 to 2018. The educational disparities of coastal for rural-urban decreased by 9.1% from 89 to 80.9. Meanwhile, educational disparities of inland for rural-urban decreased by 1.6% from 78.4 to 77.2.

On the other hand, Table 10 shows the decomposition of coastal-inland disparities for rural and urban areas. The results show that rural educational disparities of in coastal-inland provinces increased by 14.4 %, from 71.9 in 2010 to 82.3 in 2018. By contrast, urban educational disparities in coastal-inland provinces decreased by 11.2% from 85.4 to 75.8.

The results of the polarization index of educational disparities described in Table 11 calculated by equation (5) show that rural-urban educational disparities dramatically decreased during the 9-year period. This result also explains that the contribution of rural and urban educational disparities to overall educational disparities indicated a dramatic decline as described in Table 8. Most noteworthy, however, are the educational disparities for coastal-inland provinces. The results show that the polarization index of

Table 9 Inequality Decomposition by Coastal and Inland Provinces in Rural and Urban Areas (PG/NBS Data)

(unit: %)

	Rural-Urban		Rural		Urban	
	Coastal	Inland	Coastal	Inland	Coastal	Inland
2010	89	78.4	3.1	14.9	7.9	6.9
2018	80.9	77.2	5.4	30.3	13.7	10.5
Δ(%)	-9.1	-1.6	72.5	103.3	73.6	56.7

Sources: Calculated based on NBS, China Education Yearbook (2011 and 2019).

Table 10 Inequality Decomposition by Rural and Urban Areas in Coastal and Inland Provinces (PG/NBS Data)

(unit: %)

	Coastal-Inland		Coastal		Inland	
	Rural	Urban	Rural	Urban	Rural	Urban
2010	71.9	85.4	3.1	7.9	25	6.7
2018	82.3	75.8	5.4	13.7	12.3	10.5
Δ(%)	14.4	-11.2	72.5	73.6	-50.6	56.7

Sources: Calculated based on NBS, China Education Yearbook (2011 and 2019).

Table 11 Polarization Index of Educational Disparities between Coastal and Inland Provinces and Rural and Urban Areas in Each Group (PG/NBS Data)

	Rural/Urban			Coastal/Inland		
	all	Coastal	Inland	all	Rural	Urban
2010	5.14	8.09	3.64	3.69	2.56	5.86
2018	2.34	4.24	1.45	3.77	4.64	3.14

Sources: Calculated based on NBS, China Education Yearbook (2011 and 2019).

educational disparities in Table 11 increased from 2010 to 2018.

Ⅵ Comparison of Results between This Study and Previous Studies

This study has already calculated the Gini coefficient using the 2010 and 2018 aggregate data from the China Education Yearbook and 2010 and 2018 microdata from China Family Panel Studies (CFPS). It also analyzes the results described in chapter 5. In this chapter, the study compares Gini coefficient results for 2010–2018 to those for 1990–2000 (Qian and Smyth, 2008).

The results of the Gini coefficient of AYS in China as whole decreased by 20.8%, from 0.11 in 1990 to 0.0871 in 2000 (Qian and Smyth, 2008). By contrast, the results of the Gini coefficient of AYS in this study increased by 10.4%, from 0.0663 in 2010 to 0.0732 in 2018 for the 9-year period. The results of decomposition of the Gini coefficient of AYS in this study are exactly opposite those of Qian and Smyth (2008), where the results of the Gini coefficient of AYS within the coastal provinces decreased from 1990 to 2000, while those for the inland provinces decreased. In this study, however, the contribution of inequality within the coastal provinces to that of China as a whole dramatically increased by 55.7%, from 20.4% to 31.8%, in the period of 2010–2018. Educational disparities in the inland and coastal-inland provinces, meanwhile, increased from 2010 to 2018.

The results of the Gini coefficient of PG in China as whole increased from 1990 to 2000 in Qian and Smyth (2008). In this study, by contrast, the results of the Gini coefficient of PG in China as whole decreased by 38.2%, from 0.1527 in 2010 to 0.0943 in 2018. Furthermore, except for in urban areas, the Gini coefficient of PG in this study of coastal provinces, inland provinces, and rural areas decreased from 2010 to 2018. There was an increase for each group from 1990 to 2000 in Qian and Smyth (2008). Based on the given decomposition results of the Gini coefficient of PG, the contribution of the rural-urban gap to overall education inequality in this study is opposed to the previous study. In Qian and Smyth (2008), inequality within rural and urban areas decreased, but the rural-urban gap in increased by 1.3% from 82.7 in 1990 to 83.8 in 2000, within rural and urban areas it increased. In this study, within rural areas educational inequality increased by 14.7%, from 31.13 in 2010 to 35.72 in 2018, and educational inequality within urban areas dramatically increased by 68.6%, from 29.66 in 2010 to 50 in 2018. Meanwhile, the rural-urban gap decreased by 63.6%, from 39.21 in 2010 to 14.29 in 2018. On the other hand, the contribution of the coastal-inland gap to overall educational inequality in this study shared the same tendency as in Qian and Smyth (2008).

Finally, the two approaches indicate two disparities: firstly, when educational attainment is measured by AYS, the coastal-inland gap shows a slightly increasing trend; secondly, when PG is used, the gap narrows slightly. The PG approach provides more detailed information on the trend in educational disparities and provides empirical support for the argument that the rural-urban gap is the predominant contributor to overall disparities in educational attainment.

In this study, however, even if rural-urban disparities in educational attainment decreased between 2010 and 2018, the disparities that still exist between rural and urban areas do not resolve in past years. Of most note is that coastal-inland disparities in educational attainment dramatically widen over the course of 28 years.

Ⅶ Main Factors that Contribute to Educational of Education

This study has already analyzed educational disparities between coastal and inland provinces and between rural and urban areas. In order to find out what factors contribute to educational disparities in China and to what extent, I refer to the Shapley value approach based on regression analysis from Yang et al. (2014).

This study also analyzes the Gini coefficient of subgroups for each variable described in Table 12. The results of AYS and Gini for each subgroup described in Table 13 show that the Gini coefficient in China as a whole is 0.41. There dramatic disparities between rural and urban areas. The Gini coefficient for males 0.35, and for females it is 0.46. The Gini coefficient for the west is greater than for the east. Lower incomes are associated with greater disparities in the Gini coefficient. Meanwhile, younger ages are associated with smaller disparities in the Gini coefficient.

Table 14 shows disparities in the Gini coefficient associated with ***hukou***, gender, region, income, and age of AYS in 2010 and 2018. Thus, the variables described in Tables 12 and 13 are considered to affect educational attainment in this study. The model it shall estimate expresses as follows:

$$\log\ (edu) = \alpha + \beta_1 GENDER + \beta_2 HUKOU + \beta_3 AGE + \beta_4 INCOME + \beta_5 EAST + \beta_6 CENTARL \quad (15)$$

Table 12 Variable Summary (2010)

variable	N	mean	p50	sd	min	max
Gender	31,655	0.486	0	0.500	0	1
Hukou	31,655	0.295	0	0.456	0	1
Age	31,655	45.45	45	16.29	16	101
Income	31,655	10,181	6,320	16,269	50	1,000,000
Province	31,655	37.82	41	14.87	11	62
Years of School	31,655	7.014	9	4.760	0	22
Final Education	31,655	2.516	2	1.300	1	8

Sources: Calculated based on China Family Panel Studies (2010).

Table 13 Variable Summary (2018)

Variable	N	mean	p50	sd	min	max
Gender	28,197	0.499	0	0.500	0	1
Hukou	28,197	0.261	0	0.439	0	1
Age	28,197	47.41	48	16.83	16	95
Income	28,197	25,922	15,400	62,311	100	5.660e+06
Province	28,191	38.60	41	15.15	11	65
Years of School	28,191	7.660	9	5.062	0	23
Final Education	28,197	2.790	3	1.409	1	8

Sources: Calculated based on China Family Panel Studies (2018).

Table 14 AYS and Gini for Each Subgroup

		2010		2018	
		AYS	Gini	AYS	Gini
Total Sample		7.00	0.52	7.53	0.41
By Gender	Male	7.83	0.46	8.43	0.35
	Female	6.23	0.58	6.91	0.46
By Hukou	Rural	5.77	0.64	6.7	0.5
	Urban	9.97	0.24	10.42	0.14
By Age	Age ∈ [18,22]	10.26	0.28	11.83	0.18
	Age ∈ (22,26]	10.22	0.24	12.42	0.15
	Age ∈ (26,32]	9.58	0.29	11.32	0.24
	Age ∈ (32,38]	8.21	0.43	10.17	0.29
	Age ∈ (38,44]	7.11	0.53	8.71	0.38
	Age ∈ (44,50]	7.44	0.49	6.78	0.53
	Age ∈ (50,56]	6.71	0.53	6.51	0.55
	Age ∈ (56,62]	4.86	0.70	6.74	0.51
	Age ∈ (62,70]	4.55	0.73	4.4	0.70
By Income Rank	Rank = 1	4.73	0.71	4.98	0.67
	Rank = 2	6.07	0.61	6.43	0.55
	Rank = 3	6.81	0.55	7.57	0.46
	Rank = 4	7.86	0.45	8.79	0.35
	Rank = 5	9.61	0.28	10.51	0.18
By Region	East	7.69	0.47	8.13	0.35
	Central	7.34	0.49	7.85	0.38
	West	5.40	0.65	6.3	0.52

Sources: Calculated based on China Family Panel Studies (2010 and 2018).

Here, edu is the years of schooling (calculated by the variable of Years of School: 1 represents nursery school; 8 represents PHD), and income is family income. Gender (0= male and 1= female), *hukou* (0= agricultural and 1= non-agricultural), east, and central (Province represents provinces; 11 represents Beijing; 65 represents Xinjiang) are dummy variables; all their definitions are shown in Tables 12 and 13. Because the variable edu is approximately normally distributed, equation (15) can be set as a semi-log model. In addition, heteroscedasticity is a common problem in a cross-sectional data pattern, which can cause the standard error of the estimated coefficient to be ineffective. In order to test whether heteroscedasticity exists, this study adopts two testing methods to make the results robust. Table 15 demonstrates that no matter the testing approach, heteroscedasticity exists when using OLS. As a result, it is also suitable to use another method of FGLS (Feasible Generalized Least Square) designed to resolve the unprovable nature of the homoscedasticity hypothesis.

The regression results described in Table 16 show that all dependent variables are significant. Thus, the regression results can be used in an effective Shapley value decomposition.

Table 15 Heteroscedasticity Test

	2010		2018	
B-P Test (Breusch and Pagan, 1979)	223.51	(0.0000)***	6.79	(0.0092)***
White Ttest (White, 1980)	809.67	(0.0000)***	785.14	(0.0000)***

Notes: *** Significant at 1% level.
Sources: Calculated based on China Family Panel Studies (2010 and 2018).

Table 16 Regression Results of Educational Disparities

	2010				2018		
Independent Variable	Coefficient	T-Statistics	Significant Level		Coefficient	T-Statistics	Significant Level
GENDER	0.073	16.34	***		0.032	8.22	***
HUKOU	0.305	57.9	***		0.260	58.69	***
AGE	-0.009	-55.93	***		-0.011	-78.39	***
INCOME	0.000	7.43	***		0.000	4.48	***
EAST	0.049	7.58	***		0.035	6.92	***
CENTRAL	0.058	9.13	***		0.030	5.55	***
_CONSTANT	2.289	291.3	***		2.576	334.94	***
N=24195				N=19845			
R=0.2567				R=0.3174			
F-Statistic				F-Statistic			
=1272.66				=1770.25			

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$; Dependent variable is log (edu).
Sources: Calculated based on China Family Panel Studies (2010 and 2018).

Table 17 Shapley Decomposing Results of Educational Disparities

	2010	2018
Factors	Shapley Value	Shapley Value
GENDER	1.94	0.27
HUKOU	50.55	35.1
AGE	34.07	59.77
INCOME	10.79	4.07
REGION	2.65	0.8

Sources: Calculated based on China Family Panel Studies (2010 and 2018).

According to the results of Shapley value decomposition described in Table 17, the *Hukou* is the predominant factor in educational disparity in 2010. It accounts for almost 50% of educational disparity in all factors.

By contrast, age is the predominant factor in educational disparity in 2018. This paper has already discussed age as a factor in educational disparity in previous studies as a result of the Cultural Revolution (1966–76) and the lower school enrollment. Table 14 shows how individuals who are 18 to 26 years old have lower Gini coefficients than individuals who are 50 to 70 years old. On the other hand, the regression analysis results show that age as a variable is significant at the 1% level. Thus, in both the Gini coefficient and regression analysis results, age proves to be a comparatively important factor in educational attainment.

Even though the contribution of *hukou* as a contributing factor decreased from 49.12% in 2010 to 37.42 % in 2018, the Gini coefficient for rural and urban disparities in educational attainment still alerts us to a big issue in Chinese society, as discussed in chapter 3.

Age as a contributing factor to educational disparities increased by 24.58% in 2010 to 53.51% in 2018; it accounts for 53.51% of all factors in 2018. This demonstrate that age was more of a contributing factor when the *hukou* and income were less of contributing factors.

Even though the role of income as a contributing factor decreased by 19.5% to 7.78% in 2018, as this study discussed in chapter 2, income plays a crucial role in educational attainment: the *hukou* system and income accounts for over almost 40%.

Gender also plays a part in educational attainment in China. As Table 16 shows, the gender variable is significant for education, but it does not represent a major factor in the Shapley decomposing result. After the implementation of Compulsory Education in 1986, equal educational opportunities for females were widely available so that each female student could get a basic education. Even though there is still an educational gap, as evidenced by the results from Table 14, this gap stems from China's poverty-stricken rural areas, where men enjoy priority over women in receiving education as a result of a deficient supply of diversified livelihood capitals.

Finally, region does not play a major role in the Shapley decomposing result. But that does not mean that regional disparities in educational attainment no longer exist. Tables 14 and 16 show that region still impacts educational attainment as a result of uneven economic development.

Ⅷ Conclusions

This study uses NBS aggregate data and CFPS microdata, measuring regional disparities in education in China between 2010 and 2018 through the use of Gini coefficients and decomposition analysis. it locates main contributing factors to educational disparities using the Shapley decomposition approach based on regression analysis.

From the results in Chapters 5 and 6, even though the Gini coefficient of AYS in China as a whole increased from 2010 to 2018, the Gini coefficient of AYS for inland provinces decreased. The decomposition of the Gini coefficient of AYS shows that while the contribution of the coastal-inland gap to regional educational inequality as a whole decreased, within coastal provinces it increased dramatically from 2010 to 2018.

The Gini coefficient of PG in China as a whole decreased from 2010 to 2018. Except for an increase in the Gini coefficient of PG in urban areas, the Gini coefficient of PG in the remaining regional categories decreased. By decomposing the Gini coefficient of PG, the contribution of the educational inequality between rural and urban areas to overall educational inequality indicated a dramatic decline, while the contribution of inequality within rural areas and urban areas separately to overall inequality increased from 2010 to 2018. While the contribution of education inequality between the coastal and inland areas to overall inequality decreased slightly, that of inequality within the coastal provinces to overall inequality increased.

Regardless of whether we calculate the Gini coefficient of AYS using NBS aggregate data or CFPS microdata, the Gini coefficient in China as a whole increased from 2010 to 2018. However, when using these different data sets, the results of the decomposition of AYS are completely different. The results calculated by NBS aggregate data show that educational inequality within coastal provinces increased dramatically, while the educational inequality gap between coastal and inland areas decreased. By contrast, the results calculated using CFPS microdata show that educational inequality within coastal provinces dramatically decreased and the educational inequality gap between coastal and inland areas increased.

This study adopts the Shapley decomposition approach based on regression analysis to find out the factors that contribute to educational attainment disparities in China. The results show that the main factors are the age and the *hukou* system.

What can we learn from this study? As a result of economic development, access to education in China has improved over the years. A significant increase in the contribution of age to educational disparity indicates that the younger generation has more access to education than the older generation did. The results of decomposition analyses using the Gini coefficients and Shapley value show that *hukou* (rural and urban) remains a great influence on educational disparities in China, though its contribution decreased from 2010 to 2018 (see Tables 8 and 17). Table 17 indicates that the contribution of regional differences between eastern, central, and western areas to educational disparities shrunk more than anticipated over the 9-year period.

In this study, the results of regional disparities of education are different in many cases, depending on whether we use aggregate data or CFPS microdata. This implies that policy makers should examine regional educational disparities using not only aggregate data but also micro or household data.

Notes

1) The hukou system has origins in China that date back to ancient times, but the system in its current form came into being with the 1958 People's Republic of China Hukou Registration Regulation. Until very recently, each citizen was classified in an agriculture or non-agricultural hukou (commonly referred to as rural or urban) and further categorized by location of origin.

2) AYS in a given province is used to measure educational level. This study calculates the overall inequality of all 30 provinces (Tibet is excluded), then that of the 11 coastal provinces and 19 inland provinces in 2010 and 2018.

References

Angrist, J. D., and A. B. Keueger, 1991, "Dose compulsory school attendance affect schoolings?" *The Quarterly Journal of Economics,106* (4), pp. 979–1014.

Arrow, K., 1973, "The theory of discrimination," In O. Ashenfelter and A. Ress (Eds.), Discrimination in labor markets, Princeton, New Jersey: Princeton University Press.

Blanden, J., and P. Gregg, 2004, "Family income and educational attainment: a review of approaches and evidence for Britain," *Oxford Review of Economic Policy,* Vol. 20 (2), pp. 245–263.

Breusch, T. S., and A. R. Pagan, 1979, "A simple test for heteroscedasticity and random coefficient variation," *Econometrica, 47* (5), pp. 1287–1294.

Broaded, M., and C, Liu, 1996, "Family background, gender and educational attainment in urban China," *The China Quarterly, 145*, pp. 53–86.

Brown, P., and A, Park, 2002, "Education and poverty in rural China," *Economics of education review, 21* (6), pp. 523–541.

Cai, W., and F. Wu, 2019, "Influence of income disparity on child and adolescent education in China: a literature review, *Child and Adolescent Development in China. New Directions for Child and Adolescent Development, 163*, pp. 97–113.

Card, D., 1999, "The causal of education on earning," *Handbook of Labor Economics, 3*, pp. 1801–1863.

Chan, K., and L. Zhang, 1999, "The Hukou system and rural-urban migration in China: processes and changes," *The China Quarterly, 160,* pp. 818–855.

Chen, C., and C. Fan, 2016, "China's Hukou puzzle: why don't rural migrants want urban Hukou," *The China Review, 16* (3), pp. 9–39.

Cheng, T., and M. Selden, 1994, "The origins and social consequences of China's hukou system," *The China Quarterly, 139*, pp. 644–668.

Deng, Z., and D. Treiman, 1997, "The impact of the cultural revolution on trends in educational attainment in the People's Republic of China," *American Journal of Sociology, 103* (2), pp. 391–428.

Dong, Q., X. Li, H. Yang, and K. Zhang, 2008, "Gender inequality in rural education and poverty," *Chinese Sociology and Anthropology, 48* (4), pp. 64–78.

Fleisher, B. M., and X. Wang, 2004, "Skill differentials, return to schooling, and market segmentation in a transition economy: the case of Mainland China," *Journal of Development Economics, 73* (1), pp. 315–328.

Fu, Q., and Q. Ren, 2010, "Educational inequality under China's rural-urban divide: the Hukou system and return to education," *Environment and Planning A, 42* (3), pp. 592-610.

Gregg, P., and S, Machin, 2000, "The relationship between childhood experience, subsequent educational attainment and adult labor market performance," *Centre for Economic Performance, London School of Economics*, pp. 1-19.

Hannum, E., and M. Wang, 2006, "Geography and education inequality in China," *China Economic Review, 17* (3), pp. 253-265.

Jiang, S., M. Lu, and H. Sato, 2012, "Identity, inequality, and happiness: evidence from urban China," *World Development*, *40* (6), pp. 1190-1200.

Kryshak, A., 2016, "The Chinese housing registration system (Hukou): bride or wall," *Political Science and Sociology Senior Thesis,* pp. 1-31.

Lee, L., 2012, "Decomposing wage differentials between migrant workers and urban workers in urban China's labor markets," *China Economic review, 23* (2), pp. 461-470.

Ling, M., 2015, "Bad students go to vocational schools!: education, social; reproduction and migrant youth in urban China," *The China Journal, 73*, pp. 108-131.

Li, T., and J. Zhang, 1998, "Returns to education under collective and household farming in China." *Journal of Development Economics, 56* (2), pp. 307-335.

Lochner, L., and E. Moretti, 2004, "The effect of education on crime: evidence from prison inmates, arrests, and self-reports," *American Economic Review*, *94* (1), pp. 155-189.

Lucas, R. E. 1988, "On the mechanics of economic development," *Journal of Monetary Economics, 22* (1), pp. 3-42.

National Bureau of Statistical, China Eudcation Yearbook (2001, 2003, 2005, 2007, 2009, 2011, 2013, 2015, 2017, 2019) China Statitical press: Beijing.

Peng, H., L. Qi, G. Wan, B. Li, and B. Hu, 2020, "Child population, economic development and regional inequality of education resources in China," *Children and Youth Services Review, 110*, pp. 1-9.

Psacharopoulos, G., 1985, "Returns to education: a further international update and implications," *Journal of Human Resources* 20, pp. 583-604.

Psacharopoulos, G., 1994, "Returns to investment in education: a global update," *World Development 22*, pp. 1325-1343.

Qian, X., and R. Smyth, 2008, "Measuring regional inequality of education in China: widening costal-inland gap or widening rural-urban gap?" *Journal of International Development,* Vol. 20, No. 2, pp. 132-144.

Shorrocks, A. F., 2013, "Decomposition procedures for distributional analysis: a unified framework based on the Shapley value," *Journal of Economic Inequality*, *11* (1), pp. 99-126.

Thomas, V., Y. Wang, and X. Fan, 2001, "Measuring education inequality: Gini coefficients of education," vol.2525. World Bank Publications.

Wang, F., 2004, "Reformed migration control and new targeted people: China's *Hukou* System in the 2000s," *The China Quarterly*, *177*, pp. 115-132.

Wang, Y., and Y. Yao, 2003, "Sources of China's economic growth 1952-1999: incorporating human capital accumulation," *China Economic Review, 14* (1), pp. 32-52.

White, H., 1980, "A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity," *Econometrica, 48* (4), pp. 817-838.

Wu, X., and D. J. Treiman, 2004, "The household registration system and social stratification in China; 1955-1996," *Demography*, *41* (2), pp. 363-384.

Wu, X., 2011, "The household registration system and rural-urban educational inequality in contemporary China," *Chinese Sociological Review*, *44* (2), pp. 31-51.

Yang, J., X. Huang, and X. Li, 2009, "Educational inequality and income inequality: An empirical study on China," *Frontiers of Education in China*, *4* (3), pp. 413-434.

Yang, J., X. Huang, and X. Liu, 2014, "An analysis of education inequality in China," *International Journal of Education*

Development, 37 (1), pp. 2–10.

Zhang, J., and T. Li, 2002, "International inequality and convergence in educational attainment, 1960–1990," *Review of Development Economics, 6* (3), pp. 383–392.

Zhang, X., and R. Kanbur, 2001, "What difference do polarization measures make? An application to China," *Journal of Development Studies,* vol. 37 (3), pp. 85–98.

Zhao, G., J. Ye, Z. Li, and S. Xue, 2017, "How and why do Chinese urban students outperform their rural counterparts?" *China Economic Review, 45,* pp. 103–123.