The Effect of Education Expansion on Intergenerational Mobility of Education: Evidence from China

Liu, Ling and Wan, Qian

School of Economics, Xiamen University, School of Economics, Huazhong University of Science and Technology

May 2017

The Effect of Education Expansion on Intergenerational Mobility of Education: Evidence from China

Ling Liu
School of Economics, Xiamen University
Siming South Road 422# Xiamen 361005, Fujian, China
liuling7107wq@163.com

Qian Wan (Corresponding author)
School of Economics, Huazhong University of Science and Technology
Luoyu Road 1037# Wuhan 430074, Hubei, China
wqll@hust.edu.cn

Abstract

Using the data from Chinese Household Income Project, we study the effect of education expansion on intergenerational mobility of education measured with intergenerational transmission of education (ITE) through an exogenous shock, higher education expansion in 1999. Measuring ITE with years of schooling, higher education expansion (HEE) significantly decreases ITE, meaning that the gap of years of schooling between the children from different family educational background is narrowed by HEE and intergeneration mobility of education is promoted by HEE. However, when we take school quality into account and measure ITE with score of college entrance examination (CEE), HEE insignificantly decreases ITE measured with score of CEE, indicating that HEE fails to reduce the gap of higher education quality between the children from different family educational background and the inequality of higher education still maintains in some way even after HEE. We also find that ITE measured with years of schooling has an inverted-U relationship with college admission rate and ITE measured with score of CEE seems not correlate with college admission rate, which directly demonstrate the theories of MMI and EMI in the field of sociology. We further investigate the internal mechanism of the effects and we consider that the original of the inequality of higher education is the inequality of basic education. At last, we investigate the heterogeneity in the effect of HEE on ITE by gender, type of Hukou and category of CEE.

Keywords: Higher Education Expansion; Intergenerational Transmission of Education; Inequality of Opportunity
JEL classification: I21; J62

Acknowledgment: We thank that China Institute for Income Distribution provides the microdata. For useful comments, we thank Jin Xu, Sheng Fang, Jing Zhang, Haiou Mao, Shan Feng.

The Effect of Education Expansion on Intergenerational Mobility of Education: Evidence from China

1. Introduction

Education is a key mechanism for intergenerational mobility which is one of the primary topics in the study of inequality (Huang, 2013; Breen, 2010; Saar, 2010; Iyigun, 1999; Parman, 2011). Equality of education is essential for intergenerational mobility (Corak, 2013). Intergenerational mobility will increase if the education expansion decreases the gap of human capital investment between children from different family background, meaning that the newly-added educational opportunities induced by education expansion are likely to be obtained by the children from poor family background. Otherwise, intergenerational mobility will decrease.

Economists usually measure intergenerational mobility with the intergenerational correlation of socioeconomic status like income, education and occupation. Intergenerational income elasticity, denoting intergenerational income mobility, is the most common way to measure intergenerational mobility and many literatures studies the effect of educational policy on intergenerational income elasticity. Pekkarinen et al. (2009) investigate the effect of Finland's comprehensive school reform on intergenerational income elasticity, showing that this reform increases intergenerational mobility in Finland. Mayer and Loope (2008) consider that American government spending reduces the gap of human capital investment between rich and poor children and improves the intergenerational income mobility. Li et al. (2014) estimate the extent of intergenerational income elasticity in China which indicates that China's intergenerational mobility is very low and consider that government spending of public education fails to effectively promote social mobility. However, when we take the income as the measurement of socioeconomic status and measure intergenerational mobility with intergenerational income elasticity, we probably face a dilemma. Usually, we only can get the information about transitory income which obtains too much random fluctuation and measurement error. Solon (1992) shows that the extent of intergenerational income mobility is usually biased downward by measurement error, unrepresentative samples or both. In the other hand, if we take the average income of several years as permanent income to measure intergenerational income elasticity and take lifetime earnings into account, the sample size is relatively smaller, which probably restricts the expandability of the study. Precisely measuring the extent of intergenerational income mobility is a challenging task. Some economists measure the socioeconomic status with occupation, but intergenerational mobility of occupation relied on rough sorting of occupational reputation can't suitably reflect the changes in intergenerational mobility.[1] In this article, we use intergenerational mobility of education to represent intergenerational mobility and measured it with intergenerational transmission of education (ITE), which presents the effect of family educational background on children's education attainment. Education is a good measurement of socioeconomic status and has advantages relative to earnings, like less measurement error and lifecycle bias. Additionally, extensive literature has proved that higher education is associate with many other beneficial characteristics such as higher earnings, better health and longer lifespans (Black and Devereux, 2011)[2].

ITE and the effect of educational policy on ITE have been received much attention in economical literatures during the last decade. Heineck and Riphahn (2007) explore the tendency of the changes of ITE in Germany and find that the extent of ITE insignificantly decrease during the last five decades in which the German education system underwent numerous reforms in order to improve the equality of educational opportunity. Blanden and Machin (2013) claim that higher education expansion enlarges the gap of higher education attainment between rich and poor children and increases the extent of ITE in UK. Sturgis and Buscha (2015) find similar conclusions in England and Wales. Other literatures finds that some educational reforms increase ITE and reduces the relative advantages in education attainment of the children from better family educational background (Bauer and Riphahn, 2006; Bauer and Riphahn,2009). Using 1990 and 2000 Chinese Population Censuses, Magnani and Zhu (2015) employ nonparametric estimation strategies to provide a systematic investigation of ITE in urban China and find that ITE increases with time. Another strand of literatures aims at understanding the determinates or origins of ITE. Black and Devereux (2011) claim that the earliest literatures in this filed focuses on disentangling the

[1] Long and Ferrie (2013) investigate the intergenerational occupational mobility in Great Britain and the United States since 1850. They take occupations into four categories, like white collar, famer, skilled and semiskilled, and unskilled. Every category comprises many different occupation. For example, white collar comprises professional, technical, and kindred; managers, officials, and proprietors; clerical; and sales. We consider that this classification of occupation is plausible in nineteenth century because the former occupational structure is relative simply and it may face big challenge now because social division of labor is very complex in modern society.

[2] We do not mean that ITE is the best measurement of intergenerational mobility, but it seems not to be involved with measurement error or some other potential problems.

competent of intergenerational correlation due to genetics, which is predetermined and called as nature effect, and the competent due to childhood environment, which is called as nurture effect. More recently, some literature, like Black et al. (2005), Oreopoulos et al. (2006), Chevalier(2004), Maurin and Mcnally (2008), and Behrman and Rosenzweig (2002), try to identify the causal effect of parents' socioeconomic status on children's outcome. However, their conclusions are inconsistent.

Sociologists also very concern about intergenerational mobility. In the field of sociology, there are two classical theories about intergenerational mobility of education, maximally maintained inequality (MMI, Raftery and Hout, 1993) and effectively maintained inequality (EMI, Lucas, 2001). MMI implicates that the background related educational inequality first increases and then decreases and the effect of family background on education attainment will vanish when a level of education is nearly universal, so that the relationship between education supply and the extent of ITE likes an inverted-U shape. However, EMI states that social background allocates students to different types of education (school quality) and educational gap between the children from different family background always maintain even though a level of education is nearly universal. The two theories focus on different aspects of education, one is years of schooling and another is school quality. Few economic literatures takes school quality into account when investigates ITE. In this article, we focus on the following questions: How about the effects of education expansion on the ITE measured by years of schooling and the ITE about school quality? Is there a significant difference between the two effects? How about the internal mechanism of the effects of education expansion on ITE and the heterogeneity in the effects?

In 1999, China's central government made a strategic decision to expand tertiary education. From 1998 to 1999, the number of college admissions increased by 460 thousand and the college admission rate increased from 34% to 56% which represents the probability of enrolling in college. In the following ten years, this radical policy annually increased 500 thousand newly-added higher education opportunities on average, which is called a great leap forward in revolution of higher education (Li and Xing, 2010). Unlike the higher education expansion in UK which gradually increases the college enrollments, China's higher education expansion is a radical and unexpected educational reform and makes college education become universal in several years. This policy, likes a quasi-natural experiment, gives us a great opportunity to investigate the causal effect of education expansion on ITE. This paper is one of the first studies about the effect of higher education expansion (HEE) on ITE in China and contributes to understanding the effect of HEE on ITE in three ways. First, we get a comprehensive effect of HEE on ITE through measuring ITE with years of schooling and school quality. The second contribution is that we critically examine that the theory of MME and EMI. Finally, we investigate the internal mechanism of the effect of HEE on ITE and find the origins of higher education inequality.

Our main findings could be summarized as follow: When we measure ITE with years of schooling, HEE decreases ITE and ITE has an inverted-U relationship with the extent of the supply of higher education which directly demonstrates MMI. However, when we measure ITE with score of college entrance examination (CEE) denoting college quality, the empirical results show that HEE insignificantly changes ITE and ITE seems not to be correlated with college admission rate which is consistent with EMI. The internal mechanism shows that family educational background positively correlates with the type of senior high school. The marginal effect of type of senior high school on years of schooling is decreased by HEE and HEE insignificantly changes the marginal effect of type of senior high school on score of CEE, which induces the difference between the effect of HEE on ITE measured with years of schooling and the effect of HEE on ITE about school quality. At last, we find the heterogeneity in the effects by gender, type of Hukou, and the category of CEE.

The remainder of this article is organized as follow. Section 2 introduces the data set we used. Section 3 reports the empirical results about the effect of HEE on ITE measured with years of schooling. In section 4, we investigate the effect of HEE on ITE about school quality. We focus on the heterogeneity in the effects of HEE on ITE in Section 5. Section 6 is the conclusion.

2. Data

2.1 CHIP

Chinese Household Income Project (CHIP) has conducted five waves of household surveys and collects detailed information about incomes and expenditures, employment status, family structure, and social and economic characteristics at both personal and household level. HEE was implemented in 1999. The students who were firstly affected by HEE usually finished a four-years college in 2003. We only use two latest surveys in this article, CHIP 2008 and CHIP 2014. CHIP 2008 was surveyed in the early 2008 and contains 5000 households in migration sample, 8000 households in rural sample and 5000 households in urban sample. CHIP 2014 was surveyed in the

July and August 2014 and contains 7175 urban households, 11013 rural households, and only 760 migrant households. We only use the rural sample and urban sample in this article because CHIP 2014 does not nationally collect the information of migration.

2.2 Data Processing

For analyzing the effect of HEE on ITE, the main information we need is individuals' educational information and their parents'. We match individual and his or her parents via family structure information[3]. CHIP 2008 and CHIP 2014 all contain 105416 individuals and contain 89267 individuals that could theoretically match their parents' information. We successfully match 85433 individuals with their parents' information. In the matched sample, there are 70297 individuals that do not participate CEE and we drop them. Additionally, we drop the individuals who lost important variables and the individuals whose personal information is illogical, likely years of schooling is greater than age. We also drop the individuals that his/her father is sixty years older than him/her or his/her mother is fifty years older than him/her. Eventually, we get a sample with 6596 individuals for analyzing the effect of HEE on ITE measured with years of schooling and a subsample with 5760 individuals for analyzing the effect of HEE on ITE about school quality.

3. ITE measured with years of schooling

In this section, we investigate the effect of HEE on ITE measured with years of schooling and the internal mechanism of this effect. First, we conduct the empirical model for regression and investigate the effect of HEE on ITE measured with years of schooling. Second, we rule out the potential endogeneity. At last, we examine the internal mechanism. We state that we only focus on the ITE measured with years of schooling in this section.

3.1 Regress model and descriptive statistics

3.1.1 Regress model

We investigate the effect of HEE on ITE through examining that whether HEE changes the marginal effect of family educational background on children's education attainment. The linear model used to investigate the effect could be written as follow:

[3] In CHIP, every individual reports his relationship with householder, so we can precisely identify that who are the individual's parents in a family.

$$\mathrm{ed}u\_year_{ipy} = \alpha_0 + \alpha_1 Pedu_i + \alpha_2 \mathrm{e}xam_i * Pedu_i + \sum_{j=1}^{n} \beta_j \mathrm{X}_{ij} + \delta_p + \varepsilon_y + \mu_{ipy} \quad (1)$$

Dependent variable edu_year_{ipy} denotes the years of schooling of individual i who participated CEE in province p and y indicates the year in which the individual participated CEE. We take max level of parents' education, $pedu_i$, as the key independent variable denoting family educational background. $exam_i$ is a dummy variable which equals to 1 means individual i was affected by HEE, otherwise it equals to 0, and the effect of $exam_i$ on years of education is absorbed by dummies for years of CEE denoted by ε_y. δ_p represents the fix effect of the province where the individual participates the examination. X_j denotes a series of control variables, includes age, age of father and mother, dummies for gender, type of Hukou (Urban or Rural), category of CEE, and survey year. μ_{ipy} is the error term. If α_2 is significantly unequal to 0, then HEE affects ITE.

The estimator α_2 is biased if and only if variable *exam* or *pedu* correlates with error term. In fact, students can repeatedly participate CEE through repeating twelfth grade and the decision of repeating twelfth grade may correlates with some unobservable factors contained by error term. However, the time cost of repeating twelfth grade is very huge because CEE is annually hold. Particularly, no one can guarantee that he/she gets a satisfactory score of CEE through repeatedly participating CEE and the students repeating twelfth grade usually bear much psychological and emotional stress (Feng and Ding, 2007). We do not mean that no one would repeat twelfth grade for repeatedly participating CEE, but little student continuously repeats twelfth grade, so individuals can't arbitrarily choose the year of participating CEE. We consider that the endogeneity involved in this article is not serious, nonetheless, we critically take the potential endogeneity into account and rule out it in two ways.

3.1.2 Descriptive statistics

Table 1 presents the descriptive statistics of main variables. We find that the average years of schooling is 14.49 and the average level of parental education is 3.50. 55% individuals are affected by HEE, 43% are female which reflects the serious gender discrimination of human capital investment in China, and 45% Hukou were rural when they participated CEE. The average age of observations is 32.99.

Table 1 Descriptive statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Year of schooling	6596	14.49	1.969	10	24
Max level of parental education	6596	3.50	2.032	1	9
HEE (Yes=1, No=0)	6596	0.55	0.497	0	1
Gender (Female=1, Male=0)	6596	0.43	0.496	0	1
Hukou (Rural=1, Urban=0)	6596	0.45	0.498	0	1
Category of CEE (Social=1, Science=2, Art=3)	6596	1.68	0.598	1	3
Paternal age	6596	61.00	11.239	38	107
Maternal age	6596	59.38	10.927	36	98
Age	6596	32.99	8.968	17	57
Year of participating examination	6596	1998.80	9.021	1981	2012

Note: Data source is CHIP 2008 and CHIP 2014.

3.2 Empirical results

3.2.1 The effect of HEE on ITE (years of schooling)

Based on equation (1), we investigate the effect of HEE on ITE and the empirical results are presented in Table 2. In the first column, the effect of the level of parental education on children's education attainment is positive at the 1% significance level and the OLS estimator of α_1 indicates that children's years of schooling will averagely increase by 0.12 if the level of parental education increases 1. The estimated coefficient of interaction term is significantly negative which means that the gap of years of schooling between children from different family educational background is reduced by HEE when holding other variables unchanged, so HEE decreases the extent of ITE. In the next column, we narrow the time span of CEE (From 1984 to 2012) and the empirical results also show that HEE significantly lowers the extent of ITE. Continuously narrowing the time span of CEE, we get similar results in the last two columns.

Table 2 The effect of HEE on ITE (Years of schooling)

Dependent Variable: Years of schooling	ALL	1984-2012	1989-2012	1994-2012
Max level of parental education	0.119^{***}	0.131^{***}	0.142^{***}	0.160^{***}
	(0.023)	(0.022)	(0.025)	(0.034)
Max level of parental education *HEE	-0.085^{***}	-0.091^{***}	-0.087^{***}	-0.083^{***}
	(0.027)	(0.025)	(0.027)	(0.027)
Control_X	Y	Y	Y	Y
N	6596	6163	5435	4620
r2_a	0.203	0.188	0.168	0.164

Notes: The standard errors are clustered at the level of province. * p < 0.1, ** p < 0.05, *** p < 0.01.

The results in Table 2 suggest that HEE reduces ITE by 0.083 to 0.091, almost two thirds initial extent of ITE. When we measure ITE with years of schooling, we claim that HEE decreases the extent of ITE and the inequality of higher education, which means that HEE reduces the advantage in enrolling college of children from better family educational background relative to the children from poor family educational background.

3.2.2 Endogeneity

The economical and psychological cost of repeating twelfth grade are very expensive, so few students repeated twelfth grade over and over. In fact, individuals just right be affected by HEE through repeating twelfth grade if and only if they first participated CEE just a few years before 1999, therefore not all the individuals who repeat twelfth grade would induce endogeneity. If we drop the individuals who repeat twelfth grade and just right be affected by HEE, to some extent, we can solve endogeneity. We exclude the individuals who participated college entrance examination in 1999 or 2000 and the empirical results are presented in Table 3, showing that HEE significantly decreases ITE measured with years of schooling and are very similar with Table 2. We consider that the causal effect of HEE on ITE is negative and HEE increases the intergenerational mobility of education.

Table 3 Excluding the effect of repeating twelfth grade on ITE

	Dependent Variable: Years of schooling			
	ALL	1984-2012	1989-2012	1994-2012
Max level of parental education	0.120^{***}	0.132^{***}	0.143^{***}	0.163^{***}
	(0.023)	(0.022)	(0.025)	(0.034)
Max level of parental education*HEE	-0.100^{***}	-0.104^{***}	-0.096^{***}	-0.088^{**}
	(0.031)	(0.030)	(0.033)	(0.031)
Control_X	Y	Y	Y	Y
N	6116	5683	4955	4140
r2_a	0.202	0.188	0.169	0.168

Note: The standard errors are clustered at the level of province. We drop the individuals who participated CEE in 1999 and 2000. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

The correlation between birthyear and the year of participating college entrance examination is very strong. According to China educational system, we conduct a dummy variable (*Birth_dummy*) for birthyear that equals to 1 if the individual born in or after 1981 which means that the individual is expectedly affected by HEE, otherwise equals 0 which means that the individual is not expectedly affected by HEE.[4] We test the effect of birthyear on whether being affected by HHE. Table 4 presents the empirical results, showing that all estimated coefficients of *Birth_dummy* are positive at 1% significance level which proves that birthyear strongly correlates with whether being affected by HEE. More importantly, we find that family educational background has no effect on whether being affected by HEE. The estimated coefficients of other variables are also very closed to zero. In conclusion, the results in Table 4 indicate that the birthyear is the key determinant of whether being affected by HEE rather than the family educational background and unobservable factors inducing the potential endogeneity.

Table 4 The correlation between birthyear and HEE

	Dependent Variable: HEE				
	ALL	ALL	1984-2012	1989-2012	1994-2012
Birth_dummy	0.9278***	0.7379***	0.7156***	0.6928***	0.6784***
	(0.0051)	(0.0173)	(0.0199)	(0.0226)	(0.0239)
Max level of parental		0.0017	0.0021	0.0026	0.0031
		(0.0011)	(0.0012)	(0.0016)	(0.0021)
Age of father		0.0005	0.0007	0.0008	0.0010
		(0.0005)	(0.0005)	(0.0007)	(0.0009)
Age of mother		-0.0012*	-0.0015**	-0.0019**	-0.0026**
		(0.0006)	(0.0007)	(0.0008)	(0.0011)
Age		-0.0110***	-0.0135***	-0.0164***	-0.0169***
		(0.0009)	(0.0012)	(0.0017)	(0.0021)
Gender		-0.0051	-0.0068	-0.0082	-0.0095
		(0.0050)	(0.0052)	(0.0058)	(0.0065)
Hukou		0.0093	0.0132*	0.0173**	0.0200**
		(0.0063)	(0.0064)	(0.0070)	(0.0091)
Control_X	N	Y	Y	Y	Y
N	6596	6596	6163	5435	4620
r2_a	0.8571	0.8693	0.8585	0.8265	0.7304

Note: The standard errors are clustered at the level of province. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

In our sample, no one knows that China central government would implement college expansion before their birth because HEE was unexpectedly proposed in November 1998 and was rapidly implement in the next year, so variable *Birth_dummy* is an exogeneous variable. On the other hand, variable *Birth_dummy* strongly correlates with whether being affected by HEE. However, in equation (1), the possible endogenous variable is the interaction term. In practice, we take interaction term of *Birth_dummy* and *max level of parental education* as the instrument variable for the endogenous interaction term (Qian, 2008).

[4] According to China's educational system and the regulation of school admission rate, the individuals should participate CEE at 18, so we except that the individuals who born in or after 1981 participated CEE in or after 1999.

The following equation estimates the first-stage effect of interaction term constructed by dummy for birthyear and family educational background on the instrument variable constructed by dummy for HEE and family educational background.

$$exam_i * pedu_i = \alpha_0 + \alpha_1 pedu_i + \alpha_2 \mathrm{birth\_dummy}_i * pedu_i + \sum_{j=1}^{n} \beta_j X_{ij} + \delta_p + \varepsilon_y + \mu_{ipy} \quad (2)$$

The second-stage regress is as follows:

$$edu\_year_{ipy} = \lambda_0 + \lambda_1 pedu_i + \lambda_2 exam_i * pedu_i + \sum_{j=1}^{n} \beta_j X_{ij} + \delta_p + \varepsilon_y + \mu_{ipy} \quad (3)$$

Table 5 First-stage regression (ITE measured with years of schooling)

Dependent Variable: Max level of parental education*HEE				
	ALL	1984-2012	1989-2012	1994-2012
Birth_dummy*Max level of parental education	0.928***	0.926***	0.920***	0.901***
	(0.010)	(0.010)	(0.011)	(0.015)
Control_X	Y	Y	Y	Y
N	6596	6163	5435	4620
r2_a	0.928	0.923	0.913	0.891

Note: The standard errors are clustered at the level of province. * p < 0.1, ** p < 0.05, *** p < 0.01.

Table 5 presents the empirical results of first-stage regression and shows a very strong correlation between two interaction terms. Then Table 6 shows the empirical results of second-stage regression, indicating that HEE decreases ITE measured with years of schooling. The value of F-test is great than 10 which means we do not involve in the problem of weak instrument variable. Given the empirical results of ruling out endogeneity, we claim that HEE decreases ITE and improve intergenerational mobility of education.

Table 6 Second-stage regression (ITE measured with years of schooling)

Dependent Variable: Years of schooling				
	ALL	1984-2012	1989-2012	1994-2012
Max level of parental education*HEE	-0.058**	-0.065***	-0.054*	-0.042
	(0.024)	(0.025)	(0.028)	(0.037)
Max level of parental education	0.108***	0.120***	0.124***	0.131***
	(0.017)	(0.017)	(0.021)	(0.031)
Control_X	Y	Y	Y	Y
N	6596	6163	5435	4620
F_test	5011	4692.78	3731.55	1829.4
chi2	9352.008	1.8e+04	4354.079	1.6e+04
r2_a	0.205	0.189	0.170	0.166

Note: The standard errors are clustered at the level of province. * p < 0.1, ** p < 0.05, *** p < 0.01.

3.3 College admission rate and ITE (years of schooling)

The directive effect of HEE on higher education attainment is promoting college admission rate and college admission rate reflects the supply of higher education. To some extent, the effect of

HEE on ITE is similar with the effect of college admission rate on ITE. Without taking school quality into account, we analysis the relationship between college admission rate and ITE in two extreme case, in which the rate equals 0 and 1. If college admission rate equals 0, meaning that no one goes to college even though the children from the best family educational background, then ITE equals 0. In another case, college admission rate equals 1, meaning that higher education is universal and everyone goes to college even though the children from the poorest family educational background, then ITE also equals 0. Therefore, we propose that ITE first increase and then decrease while college admission rate continuously increases from 0 to 1 given the assumptions that continuous change in ITE is induced by variation of college admission rate and that ITE is always positive, which conforms with the theory of MMI.

Table 7 Test the relationship between college admission rate and ITE (Years of schooling)

Dependent Variable: Years of schooling				
	ALL	1984-2012	1989-2012	1994-2012
Max level of parental education	-0.2720^{**}	-0.3667^{***}	-0.3421^{***}	-0.7328^{***}
	(0.0946)	(0.1090)	(0.1052)	(0.1887)
Max level of parental education *	2.2828^{***}	2.7210^{***}	2.6657^{***}	4.1337^{***}
College admission rate	(0.4860)	(0.5606)	(0.5414)	(0.8255)
Max level of parental education *	-2.9030^{***}	-3.3404^{***}	-3.2756^{***}	-4.5416^{***}
College admission rate square	(0.5390)	(0.6131)	(0.5944)	(0.8248)
N	6596	6163	5435	4620
Control_X	Y	Y	Y	Y
r2_a	0.2093	0.1944	0.1759	0.1728

Note: The standard errors are clustered at the level of province. College admission rate implies the annually national college admission rate which denotes the annually-national probability of enrolling in college. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

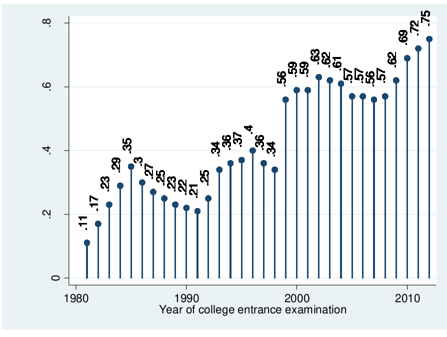

Figure1a Time trend of college admission rate

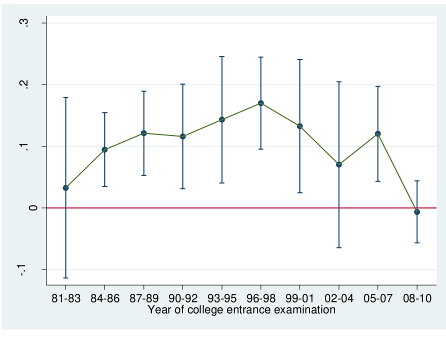

Figure1b Time trend of ITE measured with years of schooling

We consider that the relationship between college admission rate and ITE measured with years of schooling is inverted-U. We test this suspect through replacing the interaction term in equation (1) with another two interaction terms constituted by max level of parental education with college admission rate and its square and the empirical results are presented in Table 7. In Table 7, the estimated coefficients of interaction term between max level of parental education and college admission rate are significantly positive and the estimated coefficients of another interaction term between max level of parental education and square of college admission rate are significantly negative, indicating that the extent of ITE increases first and then decreases while college admission rate increases from 0 to 1 (reach peak value when college admission rate almost equals 0.4). The relationship between college admission rate and ITE measured with years of schooling is an inverted-U shape.

We examine the inverted-U relationship between college admission rate and ITE measured with years of schooling in another way. In Figure 1a, we clearly describe the time trend of college admission rate and find that college admission rate positively correlates with time trend on the whole and HEE rapidly promotes college admission rate in 1999. Based on the year of participating CEE, we take every three years individuals as a group and simultaneously estimate the extent of ITE of every group. Figure 1b describes the time trend of ITE, showing that ITE positively correlated with college admission rate before HEE and ITE negatively correlated with college admission rate after HEE, which is consistent with the empirical results in Table 2 and Table 7. In conclusion, our empirical results about the relationship between college admission rate and ITE measured with years of schooling demonstrate the theory of MMI and show that the inequality of higher education related with family background gradually vanish after HEE.

3.4 Internal mechanism

ITE represents the effect of family educational background on years of schooling. In China, whether studying in a better senior high school is very important for enrolling in a college (Park et al., 2015). First, we think that the effect of family educational background on whether studying in a better senior high school is positive. Then, we suppose that the effect of whether studying in a better senior high school on years of schooling is positive and the extent of this positive effect is decreased by HEE. Because HEE increases millions newly-added college admission places so that the students, come from a relative poor senior high school, have a greater probability of enrolling in a college after HEE. We admit that we couldn’t examine all potential mechanisms of the effect of HEE on ITE because of the limitation of data and knowledge[5].

[5] To some extent, the internal mechanism ignores the family human capital investment after the individuals enter senior high school. Due to the limitation of data, we do not investigate the difference of family human capital

Table 8 Family educational background and the type of senior high school

Dependent Variable: The level of senior high school				
	ALL	1984-2012	1989-2012	1994-2012
Max level of parental education	0.086***	0.076**	0.079***	0.048*
	(0.022)	(0.026)	(0.021)	(0.027)
Max level of parental education*HEE	-0.005	0.006	0.002	0.032
	(0.034)	(0.038)	(0.033)	(0.038)
Control_X	Y	Y	Y	Y
N	3843	3664	3360	2981
r2_a	0.059	0.056	0.057	0.054

Note: The standard errors are clustered at the level of province. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Only CHIP 2014 collects the information about the type of senior high school which the individual ever went to. We value the type of senior high school and bigger number represents better senior high school[6]. Table 8 presents the results about the correlation between family educational background and the type of senior high school. The estimated coefficients about family educational background are positive at 1% significance level, showing that children from better family educational background have a greater opportunity to study in a better senior high school relative to the children from poor family educational background. The estimated coefficients of interaction terms are insignificant which means that the relationship between family educational background and the type of senior high school keep same after HEE.

Table 9 The type of senior high school and years of schooling

Dependent Variable: Years of schooling				
	ALL	1984-2012	1989-2012	1994-2012
The level of senior high school	0.357***	0.327***	0.352***	0.366***
	(0.059)	(0.065)	(0.068)	(0.074)
The level of senior high school*HEE	-0.124**	-0.093	-0.116*	-0.128*
	(0.057)	(0.064)	(0.063)	(0.068)
Control_X	Y	Y	Y	Y
N	3843	3664	3360	2981
r2_a	0.230	0.200	0.183	0.176

Note: The standard errors are clustered at the level of province. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Next, we investigate the relationship between years of schooling and the type of senior high school. We replace family educational background with the type of senior high school which denotes

investment between the individuals in a same senior high school and we only can compare the effects of type of senior high school on years of schooling after/ before HEE, but we investigate the effect family educational background on the type of senior high school children studied which is also an important component of internal mechanism.

[6] The questionnaire take the type of senior high school into six levels. There are national or provincial key senior high school (=6), city or district key senior high school (=5), county or other level key senior high school (=4), non-key senior high school (=3), specialized secondary or vocational secondary senior high school (=2), others (=1).

in regress equation (1) and Table 9 presents the empirical results. The effect of the type of senior high school on years of schooling is positive at 1% significance level and this effect is significantly decreased by HEE which means that the gap of years of schooling between children from different type of senior high school is narrowed by HEE. Additionally, the type of senior high school positively correlates with family educational background, so the effect of family educational background on years of schooling is decreased by HEE and HEE improve intergenerational mobilty.

4. ITE about school quality

Years of schooling can't reflect the school quality. Because no matter whatever kind of four-year college you go in to, your years of schooling is 16. However, the gap of school quality between elite college and ordinary college can induce the difference of future socioeconomic status. Li et al. (2012) find that the returns to attending elite college is 10.6% even after control for student ability, major, and other factors and wage premium is larger for female and students with better-educated fathers. Based on regression discontinuity designs, Jia and Li (2016) find that the monthly wage premium of elite education is between 700-900 RMB (105-135 USD). In this section, we try to investigate the effect of HHE on ITE about school quality and get a comprehensive understanding about the effect.

4.1 Measurement of School quality

We use score of CEE to measure school quality of the college that individuals ever went to and the reasons are as follows: Firstly, higher score you get means better college you will go in to. Jia and Li (2016) show that score of CEE above the elite university cutoffs raise the admission probability by 17-19 percentage points, about 60%-70% of the mean probability. Secondly, the score also positively correlates with years of schooling. If someone wants to go to a four-years college (high-quality college), he/she usually get a higher score comparing with the individuals who only could go to a three-years college (medium-quality college). Thirdly, the college entrance examination admission system is very fair because the only one admission criteria is score of CEE.

Based on equation (1), Table 10 presents the empirical results about investigating the relationship between score of CEE and type of college.[7] In every column, we find that score of CEE

[7] Only CHIP 2013 (Urban) collects the information about the college that individuals went to. In survey questionnaire, CHIP takes the college into seven categories, 985 Project College (best college, valuing 6), 211 Project

positively correlates with type of college at 1% significance level, meaning that higher score you get and better college you will go in to. Table 11 presents the empirical results about examining the relationship between score of CEE and years of schooling, showing that score of CEE also positively correlates with years of schooling at 1% significance level. In summary, score of CEE is a feasible measurement of school quality and it also positively correlates with years of schooling.

Table 10 The relationship between score of CEE and type of college

	Dependent Variable: Type of college			
	ALL	1984-2012	1989-2012	1994-2012
Score of CEE	0.007***	0.007***	0.006***	0.006***
	(0.001)	(0.001)	(0.001)	(0.001)
Control_X	Y	Y	Y	Y
N	2518	2371	2129	1823
r2_a	0.510	0.506	0.507	0.531

Note: The standard errors are clustered at the level of province. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 11 The relationship between score of CEE and years of schooling

	Dependent Variable: Years of schooling			
	ALL	1984-2012	1989-2012	1994-2012
Score of CEE	0.008***	0.008***	0.008***	0.008***
	(0.001)	(0.001)	(0.001)	(0.001)
Control_X	Y	Y	Y	Y
N	5760	5421	4826	4141
r2_a	0.320	0.303	0.289	0.279

Note: The standard errors are clustered at the level of province. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

4.2 Regress model and descriptive statistics

4.2.1 Regress model

Based on equation (1), we replace *years of schooling* with *score of CEE* and get equation (4) to investigate the effect of HEE on ITE about school quality.

$$\mathrm{cee\_}score_{ipy} = \alpha_0 + \alpha_1 \mathrm{p}edu_i + \alpha_2 \mathrm{e}xam_i * pedu_i + \sum_{j=1}^{n} \beta_j \mathrm{X}_{ij} + \delta_p + \varepsilon_y + \mu_{ipy} \quad (4)$$

Variable cee_score_{ipy} denotes the score of CEE of individual i who participated CEE in province p and y indicates the year in which the individual participated CEE. Other variables are consistent with equation (1).

College but not 985 Project College (better college, valuing 5), other ordinary public college (good college, valuing 4), independent college (fair college, valuing 3), private college (poor college, valuing 2), adult continuing education (poorest college, valuing 1) and foreign college (best college, valuing 6). Additionally, we take only high senior school as the worst schooling quality and value it 0.

4.2.2 Descriptive statistics and analysis

Not all individuals who participated CEE report their score of CEE, the sample we used to analysis the effect of HEE in ITE about school quality is smaller than the sample we used above. Table 12 presents the descriptive statistics of the sample used to analysis the effect of HEE on ITE about school quality. We keep the individuals whose score locate between 1%-99% of the distribution. The average score of CEE is 452.75 (maximum score is 750 for most provinces).

Table 12 Descriptive statistics (school quality)

Variable	Obs.	Mean	Std. Dev.	Min	Max
Score of CEE	5760	452.75	90.286	150	650
Max level of parental education	5760	3.51	2.012	1	9
HEE (Yes=1, No=0)	5760	0.57	0.495	0	1
Gender (Female=1, Male=0)	5760	0.44	0.496	0	1
Hukou (Rural=1, Urban=0)	5760	0.46	0.499	0	1
Category of CEE (Social=1, Science=2, Art=3)	5760	1.68	0.594	1	3
Paternal age	5760	60.49	11.127	38	107
Maternal age	5760	58.91	10.837	36	98
Age	5760	32.56	8.892	17	57
Year of participating examination	5760	1999.26	8.927	1981	2012

Note: data source is CHIP 2008 and CHIP 2014.

Table 13 The difference between the two samples

	Sample A (Obs. 5760)		Sample B (Obs.836)		Analysis of diff.	
Variable	Mean	Std. Dev.	Mean	Std. Dev.	Diff.	P-Value(H_0:diff=0)
Years of schooling	14.48	1.963	14.54	1.969	-0.056	0.4430
Max level of parental education	3.51	2.012	3.41	2.16	0.097	0.1952
HEE (Yes=1, No=0)	0.57	0.495	0.42	0.493	0.156	0.0000
Gender (Female=1, Male=0)	0.44	0.496	0.41	0.493	0.023	0.2079
Hukou (Rural=1, Urban=0)	0.46	0.499	0.39	0.488	0.073	0.0001
Category of CEE (Social=1, Science=2, Art=3)	1.68	0.594	1.7	0.628	-0.017	0.4469
Paternal age	32.56	8.892	35.98	8.921	-3.429	0.0000
Maternal age	60.49	11.127	64.45	11.405	-3.96	0.0000
Age	58.91	10.837	62.61	11.006	-3.698	0.0000
Year of participating examination	1999.26	8.927	1995.68	9.053	3.58	0.0000

Note: data source is CHIP 2008 and CHIP 2014.

In Table 13, we analysis the difference between the sample of Table 12 (named it Sample A) and another sample including the individuals who do not report score of CEE (named it Sample B)[8].

[8] Sample A combines with Sample B which equals the sample in Table 1.

The last two columns show the difference between the two samples and the probability that this difference equals 0. We find that years of schooling, max level of parental education, gender and category of CEE are not significantly different in two samples. However, other variables are all significantly different in two samples at 1% significance level. We consider that recalling score of CEE is more difficult for the individuals who participated CEE earlier, which induces average age of the individuals in Sample B is greater than Sample A. Due to the difficulty of recalling score of CEE is greater than the difficulty of recalling years of schooling, the significant difference between two samples are mainly appeared in the variables associating with time. [9]

4.3 Empirical results

4.3.1 The effect of HEE on ITE (school quality)

Based on equation (4), we investigate the effect of HEE on ITE about school quality and the empirical results are presented in Table 14. The estimated coefficients of *max level of parental education* are positive at 1% significance level, showing that the level of parental education increases by 1 which induces score of CEE to increase by 3.6 to 4.1[10]. However, the estimated results of interaction terms show that HEE insignificantly decreases ITE about school quality and the inequality of higher education quality keeps the same even after HEE.

Table 14 The effect of HEE on ITE about school quality

	Dependent Variable: Score of CEE			
	ALL	1984-2012	1989-2012	1994-2012
Max level of parental education	4.017***	3.831***	3.596***	4.051***
	(0.878)	(1.032)	(1.007)	(0.821)
Max level of parental education*HEE	-0.736	-0.506	-0.079	-0.186
	(0.958)	(1.065)	(1.024)	(1.186)
Control_X	Y	Y	Y	Y
N	5760	5421	4826	4141
r2_a	0.160	0.120	0.106	0.107

Note: The standard errors are clustered at the level of province. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

4.3.2 Endogeneity

Next, we rule out the potential endogeneity using the same method mentioned above. For

[9] The degradation of memory is inevitable and there is no evidence, showing that the degradation of memory could induce endogeneity or sample selection bias.

[10] In China, CEE is the most competitive examination. Hundreds students get a same score, so you surpass hundreds competitors if your score of CEE just increases one point.

dropping the individuals who just right were affected by HEE through repeating twelfth grade, we exclude the individuals who participated CEE in 1999 or 2000 and the empirical results are presented in Table 15, showing that the empirical results in Table 14 are not involved in endogeneity induced by repeating twelfth grade.

Table 15 Excluding the effect of repeating twelfth grade on ITE (school quality)

Dependent Variable	Score of CEE				Ln(score of CEE)	
	1981-2012	1984-2012	1989-2012	1994-2012	1984-2012	1989-2012
Max level of parental education	4.100***	3.906***	3.666***	4.117***	0.013***	0.012***
	(0.882)	(1.042)	(1.013)	(0.817)	(0.003)	(0.003)
Max level of parental education*HEE	-1.146	-0.842	-0.296	-0.257	-0.004	-0.001
	(1.128)	(1.194)	(1.114)	(1.253)	(0.004)	(0.004)
N	5353	5014	4419	3734	5014	4419
Control_X	Y	Y	Y	Y	Y	Y
r2_a	0.158	0.115	0.101	0.102	0.103	0.090

Note: The standard errors are clustered at the level of province. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.We also show the empirical results when we replace dependent variable *cee_score* with its nature logarithm *ln(cee_score)*.

Table 16 First-stage regression (ITE about school quality)

	Dependent Variable: Max level of parental education* HEE			
	ALL	1984-2012	1989-2012	1994-2012
Max level of parental education* Birth_dummy	0.932***	0.930***	0.924***	0.909***
	(0.009)	(0.009)	(0.010)	(0.015)
Control_X	Y	Y	Y	Y
N	5760	5421	4826	4141
r2_a	0.930	0.926	0.916	0.895

Note: The standard errors are clustered at the level of province. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 17 Second-stage regression (ITE about school quality)

	Dependent Variable: Score of CEE			
	ALL	1984-2012	1989-2012	1994-2012
Max level of parental education*HEE	0.202	0.646	1.470	1.903
	(1.296)	(1.429)	(1.661)	(1.781)
Max level of parental education	3.720***	3.355***	2.737**	2.545**
	(0.946)	(1.117)	(1.253)	(1.132)
Control_X	Y	Y	Y	Y
N	5760	5421	4826	4141
F_test	11234.4	11488.6	9062.45	3531.54
chi2	2731.685	499.346	326.872	570.159
r2_a	0.161	0.121	0.107	0.106

Note: The standard errors are clustered at the level of province. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

As same as subsection 3.2.2, we take the interaction terms of max level of parental education and dummy for birthyear (*Birth_dummy*) as the instrument variable of interaction term in equation (4) and the first-stage regression model and second-stage regression model are similar with equation

(2) and equation (4). The empirical results of first-stage regression are presented in Table 16, showing that the instrument variable strongly correlates with endogenous variable. Table 17 presents the results of second-stage regression which indicate that the effect of HEE on ITE about school quality is insignificant.

4.4 College admission rate and ITE (school quality)

Similarly, we test the relationship between college admission rate and the ITE about schooling quality and Table 18 presents the empirical results. We do not find a significantly inverted-U relationship between college admission rate and ITE about school quality. Figure 2 also proves this and the theory of EMI is demonstrated.

Table 18 The relationship between college admission rate and ITE about school quality

	Dependent Variable: Score of CEE			
	1981-2012	1984-2012	1989-2012	1994-2012
Max level of parental education	2.5310	-1.2786	-0.4559	-0.0117
	(5.1309)	(6.8466)	(7.1995)	(0.0417)
Max level of parental education *	10.2264	26.9595	21.4906	0.1124
College admission rate	(26.2935)	(31.5809)	(32.2119)	(0.1740)
Max level of parental education *	-14.9473	-31.5105	-24.5066	-0.1184
College admission rate square	(29.3834)	(33.6331)	(33.8337)	(0.1700)
N	5760	5421	4826	4141
Control_X	Y	Y	Y	Y
r2_a	0.161	0.120	0.106	0.097

Note: The standard errors are clustered at the level of province. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

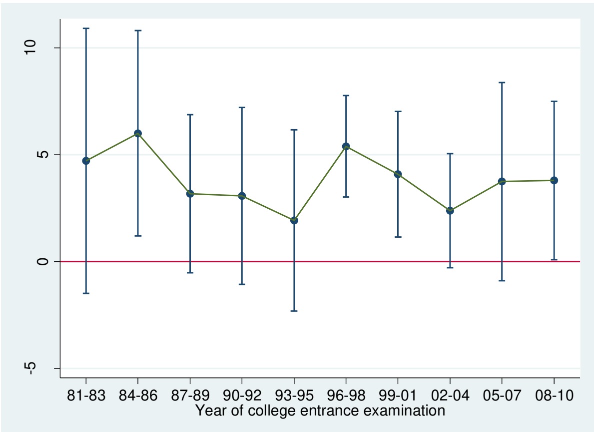

Figure 2 Time trend of ITE about school quality

4.5 Robustness check

We examine the robustness of results in Table 14 through using the natural logarithm of score of CEE and the empirical results are presented in the first two columns of Table 19, showing that

HEE insignificantly affects the extent of ITE about school quality. Furtherly, we drop the individuals who participated CEE in 1999 or 2000 and the empirical results are presented in column 3th and 4th of Table 19, showing that our conclusion seems not affected by endogeneity. At last, we investigate the effect of HEE on ITE measured with years of schooling using Sample A and the empirical results are presented in last two columns of Table 19, indicating that HEE decreases the extent of ITE measured with years of schooling and the potential problem induced by degradation of memory seems not serious. We also check the robustness of the relationship between college admission rate and ITE about school quality through using the natural logarithm of score of CEE and the time trend of ITE measured with natural logarithm of score of CEE are described in Figure 3.

Table 19 Robustness check

Dependent Variable:	Ln(score of CEE)		Ln(score of CEE)		Years of schooling	
	1984-2012	1989-2012	1984-2012	1989-2012	1984-2012	1989-2012
Max level of parental education	0.013***	0.012***	0.013***	0.012***	0.127***	0.131***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.026)	(0.028)
Max level of parental education*HEE	-0.003	-0.001	-0.004	-0.001	-0.089**	-0.079**
	(0.004)	(0.004)	(0.004)	(0.004)	(0.032)	(0.032)
Control_X	Y	Y	Y	Y	Y	Y
N	5421	4826	5014	4419	5421	4826
r2_a	0.107	0.095	0.103	0.090	0.191	0.169

Note: The standard errors are clustered at the level of province. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

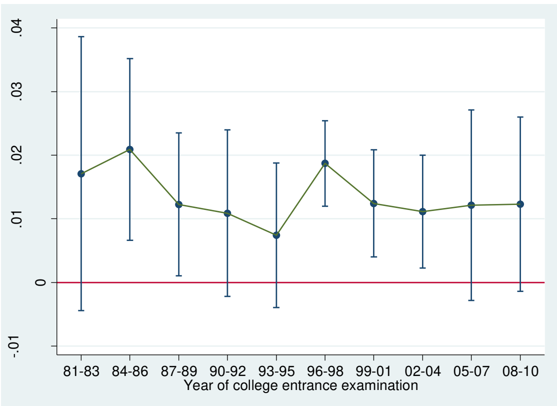

Figure 3 Time trend of ITE measured with nature logarithm of score of CEE

HEE decreases the extent of ITE measured with years of schooling and HEE insignificantly lowers ITE measured with score of CEE, so HEE reduces the inequality of higher education attainment but fails to reduce the inequality of higher educational quality. Additionally, we find that the college admission rate has an inverted-U relationship with ITE measured with years of schooling but not do correlate with ITE measured with score of CEE, which demonstrate the theories of EMI and MMI. In the other word, HEE improve the supply of higher education and higher education become nearly universal and, so HEE reduces the gap of years of schooling between children from different family educational background. However, the supply of high-quality college is always scarce and HEE is a very radical and unexpected policy, so HEE fails to significantly improve the supply of high-quality college and promote the intergenerational mobility. The inequality of higher education still maintains.

4.6 Internal mechanism

According to Table 8 and Table 9, family educational background positively correlates with the type of senior high school and HEE decreases the marginal effect of type of senior high school on years of schooling, so HEE deceases the extent of ITE measured with years of schooling. Next, we will examine that whether HEE significantly changes the marginal effect of type of senior high school on score of CEE and the empirical results are presented in Table 20. In Table 20, we find that the marginal effect of type of senior high school on score of CEE is not significantly affected by HEE, so HEE insignificantly affect the extent of ITE measured with score of CEE.

Table 20 The type of senior high school and score of CEE

	Dependent Variable: Score of CEE			
	ALL	1984-2012	1989-2012	1994-2012
The level of senior high school	17.656^{***}	17.389^{***}	17.249^{***}	19.964^{***}
	(2.181)	(2.548)	(2.796)	(2.181)
The level of senior high school*HEE	-1.365	-1.061	-0.801	-3.555
	(2.955)	(3.545)	(3.326)	(2.457)
Control_X	Y	Y	Y	Y
N	3843	3664	3360	2981
r2_a	0.221	0.180	0.161	0.160

Note: The standard errors are clustered at the level of province. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Kirst (2007) shows that the unequal educational opportunity in secondary school induces the gap of college admission rate and college quality between the children from different family background. Based on the empirical analysis about internal mechanism in section 3.4 and section 4.6, we consider that the original of the inequality of higher education is the inequality of senior high school or even the inequality in elementary education stage. Reducing the inequality of higher education not only bases on adding the supply of higher education, but also relies on lowering the inequality of basic education.

5. Heterogeneity

It is meaningful to investigate the heterogeneity in the effect of HEE on ITE. We investigate the heterogeneity in the effect of HEE on ITE by gender, Hukou, and the category of CEE.

5.1 Gender

In China, gender discrimination widely exists (Liu et al.,2000; Wie and Lee, 2017). The descriptive statistics shows that as high as 56% individuals are male and only 44% individuals are female, which indicates that male has greater opportunity to participating CEE and accept more human capital investigate than female. We investigate the heterogeneity in the effect of HEE on ITE measured with years of schooling by gender and Table 21 shows the empirical results. In Table 21, the estimated coefficients of *max level of parental education* of male is smaller than female, indicating that the intergenerational mobility of education of female is lower than male. Because of gender discrimination, family with poor economic resource tend to invest human capital for male and only the family with enough economic resource likely invest human capital for female which induce that the correlation between female human investment and family educational background is stronger than male. Further, compared with males, the effect of HEE on female ITE measured with years of schooling is greater, so the gap of years of schooling between females from different family educational background is decreased more by HEE. HEE reduces gender discrimination of educational attainment through increasing the supply of higher education.

Table 21 Heterogeneity by gender (ITE measured with years of schooling)

Dependent Variable: Years of schooling	Male		Female	
	1984-2012	1989-2012	1984-2012	1989-2012
Max level of parental education	0.097***	0.116***	0.173***	0.170***
	(0.033)	(0.036)	(0.027)	(0.024)
Max level of parental education*HEE	-0.073*	-0.076	-0.114**	-0.096**
	(0.041)	(0.043)	(0.046)	(0.040)
Control_X	Y	Y	Y	Y
N	3473	3010	2690	2425
r2_a	0.191	0.174	0.186	0.160

Note: The standard errors are clustered at the level of province. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 22 Heterogeneity by gender (ITE about school quality)

Dependent Variable: Score of CEE	Male		Female	
	1984-2012	1989-2012	1984-2012	1989-2012
Max level of parental education	3.704***	3.270**	4.054***	4.112***
	(1.199)	(1.134)	(1.129)	(1.057)
Max level of parental education*HEE	-0.057	0.660	-0.882	-0.714
	(1.496)	(1.388)	(1.599)	(1.684)
Control_X	Y	Y	Y	Y
N	3034	2658	2387	2168
r2_a	0.116	0.107	0.136	0.112

Note: The standard errors are clustered at the level of province. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 22 presents the empirical results about the heterogeneity in the effect of HEE on ITE about school quality by gender. We also find that the estimated coefficient of *max level of parental education* of female is greater than male which is also probably induced by gender discrimination. Particularly, the interaction terms are all statistically insignificant even though they are not totally equal, which means that the heterogeneity in the effect of HEE on ITE about schooling quality by gender is insignificantly. Combine with Table 21 and Table 22, HEE only reduces the gender discrimination of education attainment through decreasing the extent of ITE measured with years of schooling and does not lower the gender discrimination of school quality because HEE fails to reduce the extent of ITE about school quality.[11]

5.2 Rural and urban

The household register system, like the type of Hukou, could make the gap of opportunity to education between the children with different type of hukou (Li et al. 2015). We investigate the heterogeneity in the effect of HEE on ITE by the type of Hukou and the empirical results are presented in Table 23 and Table 24.

Based on our data, college admission rate of urban student is 49% and the rate of rural student is 29% before HEE. According to the relationship between college admission rate and ITE measured with years of schooling, the gap of college admission rate between urban and rural induces the difference of the estimated coefficients of *max level of parental education* between rural and urban in Table 23. As mentioned above, the extent of ITE reach the peak value when college admission rate almost equals 0.4, so the ITE of urban locates at the right side of the peak and the extent of ITE of rural locates at the left side of the peak. After HEE, our data shows that college admission rate of urban student is 63% and the rate of rural student is 58%, so the difference of increasing range of college admission rate between rural and urban students induces the difference of the effect of HEE on ITE measured with years of schooling between rural and urban.

[11] Lin and Lin (2012) find that Taiwan's higher education expansion does change parents' attitude on female children's education and increase the possibility for female children to attain higher education. Furtherly, parents do not change their spending behavior on children's education after the higher education expansion, which explains that the phenomenon of credentialism still prevails in Taiwanese. Credentialism is similar with pursuing elite college, so parents do not reduce the investment on their children's education which results in that intergenerational mobility of education still maintain. Unfortunately, we can't investigate the effect of HEE on family education expenditure due to the limitation of data.

In Table 24, we investigate the heterogeneity in the effect of HEE on ITE about school quality by Hukou. The difference of the estimated coefficients of *max level of parental education* between rural and urban may be induced by the gap of public education expenditure between rural and urban. The public education expenditure in rural in smaller than urban, so personal human capital investment more depend on family human capital investment in rural. However, HEE significantly increase ITE measured with score of CEE in urban and insignificantly change the extent of ITE measured with score of CEE in rural. We suppose that the family human capital investment in urban is increased by HEE which induces that the effect of family educational background on score of CEE increases. Because of the limitation of data, we can't examine this suspect.[12]

Table 23　Heterogeneity by the type of Hukou (ITE measured with years of school)

Dependent Variable: Years of schooling	Urban		Rural	
	1984-2012	1989-2012	1984-2012	1989-2012
Max level of parental education	0.106^{***}	0.113^{***}	0.163^{***}	0.177^{***}
	(0.027)	(0.031)	(0.030)	(0.041)
Max level of parental education*HEE	0.003	-0.001	-0.085	-0.096
	(0.037)	(0.038)	(0.056)	(0.063)
Control_X	Y	Y	Y	Y
N	3333	2896	2830	2539
r2_a	0.119	0.108	0.211	0.179

Note: The standard errors are clustered at the level of province. * p < 0.1, ** p < 0.05, *** p < 0.01.

Table 24　Heterogeneity by the type of Hukou (ITE about school quality)

Dependent Variable: Score of CEE	Urban		Rural	
	1984-2012	1989-2012	1984-2012	1989-2012
Max level of parental education	3.002^{**}	2.399^{**}	3.116^{*}	4.142^{*}
	(1.035)	(1.088)	(1.553)	(2.014)
Max level of parental education*HEE	2.168^{*}	2.841^{**}	0.017	-1.024
	(1.234)	(1.282)	(1.725)	(1.803)
Control_X	Y	Y	Y	Y
N	2887	2541	2534	2285
r2_a	0.121	0.111	0.121	0.108

Note: The standard errors are clustered at the level of province. * p < 0.1, ** p < 0.05, *** p < 0.01.

5.3 The category of CEE

CHIP collects the information about the category of CEE that the individuals participated. Participating different category of CEE means facing different college admission rate. Generally, comparing with the individuals who participated the CEE of social, the individuals who participated

[12] CHIP 2008 and CHIP 2014 do not collect the information about the family educational expenditure of the individuals who have already leave school.

the CEE of science and engineering face a higher college admission rate. Because China government focus on the development of nature science.

The categories of CEE are social, science and art. We only focus on the results about the categories of CEE are social and science. Because the score of CEE is not the only criterion of college admission for art students. The empirical results about this heterogeneity are presented in Table 25 and Table 26. In Table 25, The estimated coefficients of *max level of parental education* are different in two groups due to the difference of public senior high school expenditure between two groups. Compared with the students participated CEE of science, China government focus on the development of nature science and the students participating CEE of social receive less public education expenditure, so the family education expenditure is relatively more important for the students of social and the gap of human capital investment between the students of social from different family background is greater. According to our data, for the students of social, HEE increases the college admission rate by 28 percent points and for the student of science the college admission rate increases 19 percent points. The effect of HEE on college admission rate is heterogenous which induces that the heterogeneity in the effect of HEE on ITE measured with years of schooling by category of CEE.

Table 25 Heterogeneity by the category of CEE (ITE measured with years of schooling)

Dependent Variable: Years of schooling	Social		Science	
	1984-2012	1989-2012	1984-2012	1989-2012
Max level of parental education	0.162***	0.187***	0.113***	0.121***
	(0.024)	(0.028)	(0.033)	(0.033)
Max level of parental education*HEE	-0.102***	-0.114***	-0.078**	-0.067*
	(0.037)	(0.039)	(0.033)	(0.035)
Control_X	Y	Y	Y	Y
N	2393	2131	3336	2921
r2_a	0.186	0.163	0.197	0.181

Note: The standard errors are clustered at the level of province. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 26 presents the empirical results about the heterogeneity in the effect of HEE on ITE measured with score of CEE by category of CEE. The estimated coefficients of *max level of parental education* in first two columns are bigger than that in last two columns. Before HEE, most senior high schools tend to allocate more educational resource to science students which induces that the effect of family educational investment on score of CEE of science students is smaller than this effect of social students because public educational expenditure could reduce the gap of human capital investment between the children from different family educational background (Mayer and Loope, 2008). After HEE, the college admission rate of social students increases most rapidly which induces that the higher educational quality of social students decreases most rapidly, so the students from better family educational background tend to choose science and students from better family educational background tend to receive more family human investment which results in that HEE increases the ITE about school quality of science students and HEE decreases the ITE about school quality of social students.

Table 26 Heterogeneity by the category of CEE (ITE about school quality)

Dependent Variable: Score of CEE	Social		Science	
	1984-2012	1989-2012	1984-2012	1989-2012
Max level of parental education	5.631***	5.754***	2.589**	2.561**
	(1.689)	(1.491)	(1.041)	(1.020)
Max level of parental education*HEE	-3.925*	-3.985**	2.811**	3.162**
	(1.936)	(1.764)	(1.322)	(1.217)
Control_X	Y	Y	Y	Y
N	2094	1882	2960	2616
r2_a	0.126	0.106	0.103	0.097

Note: The standard errors are clustered at the level of province. * p < 0.1, ** p < 0.05, *** p < 0.01.

Table 27 The choice about category of CEE

Dependent Variable	Social (Yes=1, No=0)		Science (Yes=1, No=0)	
	1984-2012	1989-2012	1984-2012	1989-2012
Max level of parental education	-0.001	0.004	0.001	-0.003
	(0.004)	(0.004)	(0.004)	(0.004)
Max level of parental education*HEE	-0.011**	-0.017**	0.013**	0.018**
	(0.005)	(0.007)	(0.005)	(0.007)
N	5421	4826	5421	4826
Control_X	Y	Y	Y	Y
r2_a	0.063	0.066	0.069	0.072

Note: The standard errors are clustered at the level of province. * p < 0.1, ** p < 0.05, *** p < 0.01.

Table 28 The positive relationship between ITE and family educational background

	Dependent Variable: Score of CEE			
	ALL	1984-2012	1989-2012	1994-2012
Max level of parental education	1.570*	1.103	0.885	1.651
	(0.795)	(0.982)	(0.953)	(1.048)
Square of max level of parental education	0.948**	1.086***	1.118**	0.892*
	(0.335)	(0.361)	(0.393)	(0.423)
Control_X	Y	Y	Y	Y
N	5760	5421	4826	4141
r2_a	0.162	0.122	0.109	0.109

Note: The standard errors are clustered at the level of province. * p < 0.1, ** p < 0.05, *** p < 0.0.

We examine the explain that whether the students from better family educational background tend to choose CEE of science after HEE and the empirical results are presented in Table 27. The dependent variables in Table 27 are dummies for the category of CEE. According to the empirical results, the estimated coefficients of *max level of parental education* are all insignificant which means that the effect of family educational background on choosing the category of CEE is insignificant before HEE. However, after HEE, we find that the students from better family educational background relatively tend to choose science and the students from poor family educational background relatively tend to choose social. Next, we test that whether the ITE of the individuals from better family educational background is greater the ITE of the individuals from poor family educational background and the empirical results are presented in Table 28. We find that the estimated coefficients of the square of *max level of parental education* is significantly positive which means that the extent of ITE about school quality positively correlates with family educational background. Magnani and Zhu (2015) also get a similar conclusion.

6. Conclusion

In this article, we comprehensive investigate the effect of HEE on ITE. When we measure ITE with years of schooling, we find that HEE decreases ITE and ITE has an inverted-U relationship with college admission which directly demonstrates the theory of MMI. However, when we measure ITE with score of CEE denoting college quality, the empirical results show that ITE insignificantly change ITE and ITE seems be uncorrelated with college admission rate which indicates that the theory of EMI is corrected. Next, we examine the internal mechanism of this effects. We find that family educational background positively correlates with the type of senior high school. Additionally, the marginal effect of type of senior high school on years of schooling is decreased by HEE and HEE insignificantly changes the marginal effect of type of senior high school on score of CEE. At last, we investigate the heterogeneity in the effect of HEE on ITE by gender, type of Hukou, and category of CEE. When we explain the heterogeneity in the effect of HEE on ITE measured with years of schooling, we often focus on the difference in higher education supply. However, when we explain the heterogeneity in the effect of HEE on ITE about school quality, we often focus on the difference in higher education demand or human capital investment. Unfortunately, due to the limitation of data, we can’t examine the effect of HEE on family human capital investment. We conclude that HEE narrow the gap of years of schooling between the children from different family educational background but the gap of human capital between the children from different family educational background still maintain in some way.

High quality educational resource, especially for higher education, is very scarce. Education expansion may lower the gap of years of schooling between the children from different family background, but the inequality of education still maintains in some way. According to the internal mechanism discussed in this article, we claim that the origin of the inequality of higher education is not the higher educational system nor the college admission system rather than the inequality of basic education due to the school district system and the inequality of public educational expenditure between different school districts.

References

Behrman J R, & Rosenzweig M R. (2005). Does Increasing Women's Schooling Raise the Schooling of the Next Generation. American Economic Review, 95(5):1738-1744.

Black, S. E. & Devereux, P. J. (2011) Recent developments in intergenerational mobility, in Handbook of Labor Economics, Vol. 4, Ashenfelter,O. and Card,D.(Eds), North Holland, Amsterdam, pp. 1487–541.

Blanden, J., & Machin, S. (2013). Educational inequality and the expansion of uk higher education. Scottish Journal of Political Economy, 60(5), 230-249.

Bauer, P., & Riphahn, R. T. (2006). Timing of school tracking as a determinant of intergenerational transmission of education. Economics Letters, 91(1), 90-97.

Bauer, P. C., & Riphahn, R. T. (2009). Age at school entry and intergenerational educational mobility. Economics Letters, 103(2), 87-90.

Breen, R. (2010). Educational Expansion and Social Mobility in the 20 th Century. Social Forces, 89(2), 365-388.

Chevalier, Arnaud, (2004). Parental education and child’s education: a natural experiment. Discussion Paper No. 1153, Institute for the Study of Labor (IZA).

Corak, M. (2013). Income Inequality, Equality of Opportunity, and Intergenerational Mobility. The Journal of Economic Perspectives, 27(3), 79-102.

Feng Y., & Ding L. (2007). Investigation of the Mental Health Status of 174 Returnees from Senior High School Graduate before the College Entrance Examination. China Journal of Health

Psychology, 15(5), 389-392.

Heineck G. & Riphahn R T., 2007, Intergenerational Transmission of Educational Attainment in Germany: The Last Five Decades. Discussion Paper No. 2985, Institute for the Study of Labor (IZA).

Huang, J. (2013). Intergenerational transmission of educational attainment: the role of household assets. Economics of Education Review, 33(74), 112–123.

Iyigun, M. (1999). Public Education and Intergenerational Economic Mobility. International Economic Review, 40(3), 697-710.

Jia R., & Li H. (2016). Access to Elite Education, Wage Premium and Social Mobility: The Truth and Illusion of China's College Entrance Exam. Stanford Center for International Development, Working Paper NO.577.

Kirst, M. (2007). Secondary and Postsecondary Linkages. In Dickert-Conlin S. & Rubenstein R. (Eds.), Economic Inequality and Higher Education: Access, Persistence, and Success (pp. 44-66). Russell Sage Foundation.

Li, H., Meng, L., Shi, X. & Wu, B. (2012). Does Attending Elite Colleges Pay in China?" Journal of Comparative Economics, 40(1), pp. 78 − 88.

Li, H., Loyalka, P., Rozelle, S., Wu, B. & Xie, J. (2015)."Unequal Access to College in China: How Far Have Poor, Rural Students Been Left Behind?, The China Quarterly, Vol.221, 185–207.

Li, Shi, & Xing, C. (2010). China's higher education expansion and its labor market consequences. Social Science Electronic Publishing.

Li, Z., Liu, L., & Wang, M. (2014). Intergenerational income mobility and public education spending: evidence from china. Children & Youth Services Review, 40(3), 89-97.

Lin, C., & Lin, C. (2012). Does Higher Education Expansion Reduce Credentialism and Gender Discrimination in Education? Social Indicators Research, 109(2), 279-293.

Liu, P., Meng, X., & Zhang, J. (2000). Sectoral Gender Wage Differentials and Discrimination in the Transitional Chinese Economy. Journal of Population Economics, 13(2), 331-352.

Long, J., & Ferrie, J. (2013). Intergenerational Occupational Mobility in Great Britain and the United States Since 1850. The American Economic Review, 103(4), 1109-1137.

Lucas, S. (2001). Effectively Maintained Inequality: Education Transitions, Track Mobility, and Social Background Effects. American Journal of Sociology, 106(6).

Magnani, E., & Zhu, R. (2015). Social mobility and inequality in urban china: understanding the role of intergenerational transmission of education. Applied Economics, 47(43), 1-17.

Mayer, S. E., & Lopoo, L. M. (2008). Government spending and intergenerational mobility. Journal of Public Economics, 92(1–2), 139-158.

Maurin, E., & McNally, S. (2008). Vive la Révolution! Long-Term Educational Returns of 1968 to the Angry Students. Journal of Labor Economics, 26(1)

Oreopoulos, P., Page, M., & Stevens, A. (2006). The Intergenerational Effects of Compulsory Schooling. Journal of Labor Economics, 24(4), 729-760

Park, A., Shi, X., Hsieh, C. T., & An, X. (2015). Magnet high schools and academic performance in china: a regression discontinuity design. Journal of Comparative Economics, 43(4), 825-843.

Parman, J. (2011). American Mobility and the Expansion of Public Education. The Journal of Economic History, 71(1), 105-132.

Pekkarinen, Tuomas, R. Uusitalo, and S. Kerr. (2009). School tracking and intergenerational income mobility: Evidence from the Finnish comprehensive school reform. Journal of Public Economics NO.93.7-8:965-973.

Qian, N. (2008). Missing Women and the Price of Tea in China: The Effect of Sex-Specific Earnings on Sex Imbalance. The Quarterly Journal of Economics, 123(3), 1251-1285.

Raftery, A., & Hout, M. (1993). Maximally Maintained Inequality: Expansion, Reform, and Opportunity in Irish Education, 1921-75. Sociology of Education, 66(1), 41-62.

Saar, E. (2010). Changes in Intergenerational Mobility and Educational Inequality in Estonia: Comparative Analysis of Cohorts Born between 1930 and 1974. European Sociological Review, 26(3), 367-383.

Solon, G. (1992) "Intergenerational Income Mobility in the United States." The American Economic Review, Vol. 82, No. 3, pp. 393–408.

Sturgis, P., & Buscha, F. (2015). Increasing inter-generational social mobility: is educational expansion the answer?. The British Journal of Sociology, 66(3), 512-533.

Wie, D., & Lee, J. W. (2016). Wage structure and gender earnings differentials in China and India. World Development, In Press, Available online 8 May 2017.

Acknowledgement

We thank that China Institute for Income Distribution provide the microdata in this article. For useful comments, we thank Jin Xu, Sheng Fang.