HHS Public Access

Author manuscript

Comp Educ Rev. Author manuscript; available in PMC 2015 August 01.

Published in final edited form as:

Comp Educ Rev. 2014 August ; 58(3): 509–536. doi:10.1086/676828.

Two Aspects of the Rural-Urban Divide and Educational Stratification in China: A Trajectory Analysis*

Lingxin Hao,

Department of Sociology, Johns Hopkins University, 3400 N. Charles Street, Baltimore, MD 21218 U.S.A., Tel: (410) 516-4022, Fax: (410) 516-7590

Alfred Hu, and

Department of Sociology, Fu-jen Catholic University, Taiwan, Tel: 886-02-29052197

Jamie Lo

Department of Sociology, National Tsing Hua University, Taiwan, Tel: 886-911105411

Lingxin Hao: hao@jhu.edu; Alfred Hu: alfred.kw.hu@gmail.com; Jamie Lo: m4222@msn.com

Abstract

Contextualized in China's social change of the past half-century, this paper develops the notion of dichotomous inequality to conceptualize the two aspects of China's rural-urban divide in educational inequality—the household registration system (*hukou*) assigns people to a top-bottom hierarchy, and the rural-urban schooling system institutionalizes unequal resource distribution and diverse school mission. Based on this conceptualization, we formulate a Chinese version of the maximally maintained inequality (MMI) hypothesis. We capitalize on individual educational history data from the China General Social Survey (CGSS) 2008 and conduct a trajectory analysis using the generalized mixture modeling to estimate the differential effects of the two aspects of rural-urban divide on educational inequality in China. Findings indicate that (1) the sorting mechanism of the rural *hukou* places rural-*hukou* people in the very bottom of educational stratification, (2) the penalty of attending rural pre-tertiary school increases with educational stages, and (3) there is a cumulative disadvantage of rural *hukou* and rural school. Overall, our findings attest to the Chinese-version MMI and the behind principle of inequality reproduction.

Keywords

rural-urban divide; *hukou*; educational system; educational stratification; China

Introduction

In China, the rural-urban divide has been considered the main driver of educational stratification (Hannum 1999; Knight and Shi 1996; Wu and Treiman 2004). Some scholars conceptualize the rural-urban divide to be the rural-urban household registration (*hukou*) dichotomy (e.g., Wu and Treiman 2004), while others focus on educational policies regarding resource allocations and educational expansions that contribute to rural-urban

*The paper was presented at the RC28 Spring meeting, Hong Kong, May, 2012.

differences in educational provision, accessibility, and quality (Hannum 1999; Zhang and Kanbur 2005). Previous research has examined one but not both aspects of the rural-urban divide in shaping social outcomes. For example, researchers have found *hukou* to determine educational attainment (Zhou et al. 1998), educational transition (Wu 2010), and intergenerational occupational mobility (Wu and Treiman 2007); rural-urban differences in educational systems has explained inequality in school enrollment and educational attainment (Hannum et al. 2007). Little is known, however, about the relative importance and differential roles of the two aspects of the rural-urban divide in maintaining and re-engendering educational inequality.

The two aspects of the rural-urban divide are related but conceptually distinct. The *hukou* system defines two types of citizenship in a single country, where rural-*hukou* citizens are regarded as secondary to urban-*hukou* citizens in citizen rights, such that rural-*hukou* citizens are restricted from accessing public goods available in urban areas (Chan 1994). Distinctively, the educational system defines schools as either rural or urban, and rural schools are secondary to urban schools in resource allocation (Fu 2005). Compared to urban schools, rural schools are poorer in financial funding, teacher qualifications, and educational quality and, consequently, they have lower levels of school enrollment, student aspiration, and student achievement.

However, the boundaries of the two aspects of the rural-urban divide—rural-urban *hukou* and rural-urban school—do not completely overlap, given two facts. First, educational policies are implemented through the educational bureaucracy, and the lowest administration for primary and secondary education is the county department of education for rural schools and the city/district department of education for urban schools (Ministry of Education of China 2012). Since some residents in townships have urban *hukou*, schools in townships serve both rural-*hukou* and urban-*hukou* populations. Thus, when urban-*hukou* children in townships go to local schools, as most of them do, they attend schools administered under the county department as rural schools. Second, the *hukou* system has gradually relaxed to allow—without changing their *hukou* status—geographic mobility of rural-*hukou* people, which has been increasing since 1978 throughout the reform era, and rural-*hukou* children brought along by their migrant parents attend urban schools (Author et al. 2012). These variations in exposure to rural-urban schools within *hukou* groups allow us to investigate the two distinct aspects of the rural-urban divide simultaneously.

To address the two aspects of the rural-urban divide in educational stratification, we take a trajectory perspective, which views the patterns of educational behaviors and events evolving over an individual's life course (Elder 1985). With individuals' educational trajectories constructed on the basis of educational histories from the China General Social Survey (CGSS) 2008, we examine educational stratification from 1949, the year when the People's Republic of China was founded, to 2008, the year of deepening marketization. It is under the broad social changes in the past half-century that we ask how the two aspects of the rural-urban divide—rural-urban *hukou* and the rural-urban schools—maintain or reshape inequality in educational attainment.

Educational Stratification in China

The *hukou* dichotomy and the rural-urban school dichotomy are major drivers of educational inequality in China. In the 1950s, the *hukou* dichotomy was used to assign people to agricultural (rural) *hukou* vs. nonagricultural (urban) *hukou* status, initially according to one's place of residence and then according to one's mother's *hukou* status, creating a new social hierarchy (Chan 1994). People assigned rural *hukou* were placed at the bottom layer of the new hierarchy, paying high agricultural taxes while having limited access to public goods in all life domains, including education, medical care, and retirement pensions, all of which were readily available to urban-*hukou* people (Whyte and Parish 1985). As an integral part of China's urban bias, the *hukou* dichotomy has been relatively stable despite the economic reform since 1978 and the recent relaxation of geographic mobility that allows rural residents to work in cities. Most rural migrants, however, are not permitted to convert their rural *hukou* to urban *hukou* status(Cheng and Selden 1994; Chan 2009). Nowadays, *hukou* is still the *de facto* identity card used to distinguish between the top and bottom layers of the *hukou* hierarchy (Chan 2009).

The rural-urban school system stratifies educational opportunities and outcomes at the institutional level, different from the structural force of *hukou*. Rural schools have been historically and contemporarily inferior to urban schools. Educational development in the vast rural areas has lagged behind that in cities (Lee 2006). Like most socialist states, which are urban centered, the Chinese central government favors cities over counties when allocating resources (Szelenyi 1983; Titma and Tuma 1993; Zhou et al. 1998). Schools in central-controlled municipalities, such as Beijing and Shanghai, followed by provincial capitals, enjoy greater resources from the central government and thus enjoy higher school quality. In contrast, most rural schools are locally funded and thus of lower school quality. The unequal resource allocation under pre-reform centralized educational policies contributed to the resource gap between rural and urban school, and this gap has been growing since the policy shifted to a liberal competitive model of education after the economic reform began in 1978 (Hannum and Xie 1994; Hannum 1999; Pepper 1980). In particular, a shift toward decentralization of educational finance policy in the early 1980s reduced resource redistribution to rural areas (Tsang 2000). Because rural areas are poorer than urban areas, decentralization has exacerbated the financial constraints on rural schools. This has deepened what scholars have called "educational stratification by geography" (Hannum and Wang 2006; Lee 2008). In effect, the decentralization of education maintains the rural-urban gaps in the school system, despite the continuous educational expansion since the enactment of the 1986 compulsory education law.

The two aspects of rural-urban divide have evolved very differently over the past half-century. The *hukou* dichotomy has been extremely stable, weathering various political campaigns, the Cultural Revolution, the economic reform, and the current deepening marketization of the economy. In contrast, educational policies have let educational systems ride with the social-change tide, going through early reconstruction and expansion, suspension during Cultural Revolution, decentralization and marketization in the 1980s, and rapid expansions of higher education in the last decade. The fixed nature of the *hukou* dichotomy and the fluidity of educational policies each uniquely contribute to the rural-

urban inequality in educational opportunities and outcomes across a wide range of birth cohorts. Teasing out the effects of the two aspects thus provides an excellent opportunity to advance our understanding of China's educational inequality.

Clearly, the two aspects of rural-urban divide, *hukou* and school, stratify educational opportunities and outcomes differently. The literature to date, however, fails to identify how these two dichotomies affect educational inequality differentially. The estimate for one aspect of the rural-urban divide in this previous literature captures the *confounding* effect of both aspects, thereby compromising conceptual clarity. The lack of longitudinal data and the insufficient variation in the two aspects of rural-urban divide in cross-sectional data are the major sources of this limitation. For example, Wu and Treiman (2007) recognized the difference between the *hukou* and geographic aspects of rural-urban divide but analyzed only *hukou* and dropped place of residence partly because the two variables overlapped to a large degree in their cross-sectional data. Thus, decoupling the effects of the two aspects through use of longitudinal data would yield a correct understanding of the mechanisms by which inequality is maintained and re-engendered, so as to provide viable policy recommendations for reforming the *hukou* system and the educational system.

This paper exploits event-history-calendar data on individuals' educational experiences in the CGSS 2008 to construct longitudinal data for 5,187 sample individuals. The longitudinal data help identify individuals' *hukou* status assigned at birth, childhood *hukou* change, and pre-tertiary school location, which do not overlap completely and provide sufficient variation for estimating their respective partial effects on educational trajectories.

Theory and Research of Educational Stratification

Contemporary educational stratification is studied in the context of educational expansions, because such expansions are nearly universal throughout the world (Boli et al. 1985). How educational stratification changes during mass education expansions has generated several strands of explanations. The stratification of education by macro factors has been theorized under three schools of thought. The *modernization hypothesis* predicts that the effect of family origins will decrease with the progress of modernization (Lipset 1959). The *social reproduction hypothesis* expects that the decrease in origin effects on lower educational transitions is compensated for by increasing effects on later transitions (Bowels and Gintis 1976). The *socialist transformation hypothesis* anticipates that the initial reduction in origin effects will turn around as new elites pursue their interests (Shavit and Blossfeld 1993).

Among the numerous empirical studies on educational stratification, Mare (1981) suggested that the allocation and distribution of schooling are conceptually independent and may change in response to demographic and behavioral changes. Mare's work won popularity in empirical analyses of schooling continuation decisions and educational transitions to illuminate how educational expansions do or do not lead to the reduced effect of family origins on educational inequality in complicated ways. In a widely cited comparative study of educational transitions directed by Shavit and Blossfeld (1993), global expansion in mass education provided a basis for comparison across countries at different developmental stages and under different political authorities. Based on this comparative study of educational

stratification in 13 industrialized countries in Asia, America, and Europe, Shavit and Blossfeld explored which societal conditions help reduce existing inequalities in educational opportunities. They found that, in most industrialized countries, inequalities in educational opportunities among students from different social and economic strata had been remarkably stable since the early 20th century. This finding offers little evidence to support the modernization hypothesis, except for evidence from Sweden and the Netherlands, where the origin effects weakened (see Breen and Jonsson's [2005] review).

Based on the Irish educational stratification during an educational expansion, an additional hypothesis (referred to as MMI) posits that educational inequality can be maximally maintained during educational expansions (Raftery and Hout 1993). The Irish experience suggests that origin effects decline when privileged groups' participation reaches "saturation" (full participation) at a lower level of education; at the expanded higher level, though, saturation is absent and origin effects are maintained, even though the expansion pushed up the overall educational attainment level. The MMI is consistent with the *social reproduction hypothesis* in that reduced inequality at a lower level of education is possible, but this reduction is offset by the enlarged inequality in transitioning to higher levels of education. What makes the MMI unique is its clear identification of a necessary condition under which a reduction of inequality is possible—saturation of all members of the privileged groups—and the important insight into why an increase in inequality under expansion is equally possible when saturation of the privileged groups at the highest level has not been reached. The condition of saturation is quantifiable and thus empirically testable.

An important critique of the MMI is that it ignores the quality of education (Breen and Jonsson 2000; Lucas 2001). According to Lucas (2001), tracking, *de jure* or *de facto*, in U.S. schools places racial minority students in lower educational tracks and effectively maintains educational inequality at all educational levels, independent of quantitative equality. This is called the *effectively maintained inequality hypothesis* (EMI). Testing the EMI requires information on educational quality and especially school tracking, which are not available in data on the adult population like the data of this paper.

A third addition to the explanations about education stratification is the resurgence of rational choice models (RCM) focusing on educational decision making (Erikson and Jonsson 1996; Breen and Goldthorpe 1997; Morgan 1998). Adopting the *utility maximization perspective* from economists (Cameron and Heckman 1998), students and their parents consider the expected benefits, costs, and probability of success for educational alternatives in their decision making. Decision-making models are required to test this hypothesis appropriately.

When looking at individuals' educational trajectory, the *cumulative advantage thesis* reviewed in DiPrete and Eirich (2006) offers further insights. The cumulative nature of educational attainment, i.e., that a higher-level attainment is built on a lower-level attainment, suggests that the advantage of the privileged social groups at the beginning of an individual's educational trajectory will compound throughout educational transitions. The mechanisms include the compound returns to earlier human capital stock and the access to

increasingly greater resources brought about by earlier higher achievement. Likewise, the disadvantage of lower social groups at the beginning of the trajectory will compound. Thus, initial social positions are driving forces of divergent educational trajectories.

As a socialist country in transformation, China has not yet been included in cross-national comparative studies. Since China's educational expansion began in 1986, much empirical work has focused on the changing effects of family origin on educational inequality. Addressing the changing inequality before and after the 1999 higher education expansion, Li (2010) considered that China's social change might condition the family-origin effects during a specific time of educational expansion in different ways from those for the industrialized societies in which the MMI, EMI, and RCM were developed. Li proposed that China's peculiarity in the marketization of tertiary education went hand-in-hand with its fivefold expansion, resulting in soaring costs of college education. Using micro-data from the 2005 China Census, the author tested MMI, EMI, and RCM. While the findings largely supported the hypotheses, the author highlighted the different finding for the rural-urban origin effect, which was increased rather than preserved after the 1999 higher education expansion, as had been the case for other origin effects. Similarly, Guo (2010) found the rural-urban gap in transition to senior-high schooling declined, while the gap in the transition to college increased after higher education was expanded. Both studies point to the need to further investigate educational stratification according to the rural-urban divide. This paper fulfills that need by conceptualizing and estimating the role of the two aspects of rural-urban divide in maintaining and reshaping educational inequality.

Conceptual Framework and Hypotheses

Our study of educational stratification focuses on the rural-urban divide in the past half-century under China's socialist regime. China's rural-urban divide in educational opportunities and outcomes can be fundamentally conceived as what Charles Tilly (1998) called *categorical inequality*, a pattern of social inequality that is remarkably "durable." Three powerful mechanisms make categorical inequality possible. The first mechanism assigns people to social categories; the second mechanism institutionalizes the practices that allocate resources unequally across these categories (Massey 2007). In the third mechanism, advantages are cumulative (e.g., advantages beget advantages) throughout the educational stages (DiPrete and Eirich 2006).

For Tilly (1998), the unequal allocation of resources occurs through exploitation, where one group expropriates resources produced by another group; and through opportunity hoarding, where one group restricts another group's members' access to scarce resources or excludes them socially. This unequal allocation is durable over time through emulation, whereby mechanisms of exploitation and exclusion are copied throughout various life domains and across time periods; and through adaptation, whereby people begin to accept and behave according to the social rankings. These ideas are perfectly exemplified in China's rural-urban divide. In our view, China's educational stratification by the rural-urban divide is a typical phenomenon of categorical inequality, or more precisely, dichotomous inequality.

Massey (2007) explains the mechanisms by which the structural inequality is translated to the micro level through institutions assigning people to hierarchical social groups and allocating resources unequally among these groups. In China, the *hukou* system assigns people to rural *hukou* (the lower layer) vs. urban-*hukou* (the upper layer) according to the mother's *hukou* status. The opportunity hoarding has been exemplified by the *hukou* dichotomy where rural-*hukou* people's access to opportunities in multiple life domains are limited compared to those available for urban-*hukou* people, such as education, employment, healthcare, and pensions. In addition, high agricultural taxes were a heavy burden on the rural population. The Chinese rural-urban educational system has historically and contemporarily been segregated under an enormous development-level gap. The rural-urban school missions were distinct, despite educational expansion: the mission of rural schools was literacy before and compulsory education now; the mission of urban schools was compulsory education before and college-oriented now. Educational policies before and after the college expansion allocate more resources to urban schools than rural schools. Thus, the rural primary school a person attended confers a positional disadvantage for entry into a high-quality junior high school, which is usually an urban school. This chain continues on to the highest level of education, resulting in a tiny probability for this rural person to attend college.

The cumulative advantage thesis brings the dynamics of inequality to the front. Over time, exposures to a "treatment", be it minority status, family poverty, low school quality, or low-level academic tracking, may result in cumulative disadvantage (Blau and Duncan 1967; Duncan et al 1998; McLanahan and Sanderfur 1994; Lucas 1999). In China, limited vs. unlimited access to resources by *hukou* will compound over time, leading to divergent outcomes, such as educational attainment. Exposure to rural pre-tertiary school is a disadvantage, which also compounds through educational transitions. In all, the assignment to rural or urban *hukou* group, the exposure to rural or urban primary and secondary school, and the compounding disadvantage educational process are mechanisms linking the two aspects of rural-urban divide to individual educational attainment in China.

Our conceptual framework modifies the MMI in a dynamic perspective. First, the saturation condition of the MMI refers to a privileged family origin as measured by parental socioeconomic status or SES. In developed countries, SES groups are stable in that higher SES are always more privileged and lower SES are always less privileged. In contrast, Chinese status groups have been extremely unstable, given the reverse class structure and a new definition of "class lines" (pre-revolution landlords, rich farmers, counterrevolutionaries, and intellectual rightists) after the revolution, the ups and downs of high-powered officials during the Cultural Revolution, the elimination of the class lines in 1979, and the rise of new elites combining wealth and political power since the economic reform. All this makes the role of family SES change over time or even over individual educational trajectories. However, the dichotomous rural-urban divide in China is durable, with the *hukou* dichotomy labeling people in a fixed way throughout the history of socialist China and over individuals' life courses. Thus, a corresponding test of the MMI regarding family origins in China is to examine the *hukou* effect.

Second, while industrialized societies expand their education independent of social-group boundaries, the Chinese case exhibits a peculiar exception in that educational expansions have different implications for rural vs. urban schools. For example, when the expansion targets higher education, rural schools are irrelevant, because their mission was merely moving from literacy to compulsory education. With expanded college education, urban pre-tertiary schools are quickly adapting to update their mission to be college-oriented. These divergent impacts of educational expansion on rural vs. urban educational systems have profound consequences for individuals' educational trajectories. Rural pre-tertiary schools expose students to the institutionalized resource restrictions, low mission, and social exclusion, leading to lower educational aspiration, lower school enrollment, greater likelihood of schooling termination, and lower educational attainment. Thus, to adequately test the MMI in China's case, one must examine the rural-urban school dichotomy.

Third, both *hukou* dichotomy and school location dichotomy lead to the accumulation of advantage or disadvantage over individuals' educational career. The compounding effect of childhood exposure to rural *hukou* and rural school is the slow growth of years of schooling, whereas that of exposure to urban *hukou* and urban school is the steady growth of years of schooling. A trajectory view of individuals' educational history is excellent to capture cumulative advantage and disadvantage. Exposure to rural *hukou* and rural school will be likely to sort individuals to low schooling trajectories; in contrast, exposure to urban *hukou* and urban school will be likely to sort individuals to high schooling trajectories. Thus, to test cumulative advantage/disadvantage under MMI, we must examine whether there are divergent educational trajectories driven by the two aspects of rural-urban divide and whether their effects are stronger for later educational stages.

Based on these rationales, we derive three testable hypotheses. First, the fixation of rural-urban *hukou* dichotomy and the very bottom position of rural-*hukou* people in multiple life domains suggest that rural-*hukou* people are more likely to follow the lowest educational trajectory (*Hypothesis 1*). Second, because pre-tertiary schools institutionalize unequal resource allocation and different school missions with a more profound impact on transitioning to above-compulsory schooling than completing below-compulsory schooling, we expect that rural pre-tertiary schools lower their students' probability of following the highest trajectory (*Hypothesis 2*). The notable difference between the *hukou* effect in Hypothesis 1 and the school effect in Hypothesis 2 is the educational stratification—the bottom position for the *hukou* effect and the top position for the school effect. Third, the disadvantage of rural *hukou* and rural school not only sorts people into distinct educational trajectories, it also accumulates over individuals' educational career. In other words, the disadvantage is greater for later educational stages than earlier ones (*Hypothesis 3*).

Data and Methods

Much educational stratification research has been conducted on the reduced-form effects of social stratification factors on ultimate formal educational attainment (Blau and Duncan 1967) and the temporal stability of these effects (Hauser and Featherman 1976). Since 1980 this research literature has been dominated by the school continuation and school transition models (Mare 1980, 1981). Researchers have disaggregated formal schooling attainment

into a series of grade transitions and analyzed the variation over transitions when looking at the effects of social stratification factors on school continuation. Although earlier research missed out on examining the educational process, later research missed out on conveying an overall picture of educational stratification. This paper applies generalized mixture modeling (GMM) (Muthén and Sudden 1999; Muthén 2004) to analyze individuals' educational histories. The GMM trajectory analysis has been increasingly adopted to address a variety of issues, such as those related to child development, mental health, and academic achievement (author 2012; McLeod and Fettes 2007; Stoolmiller et al. 2005). Our purpose in using GMM is to address the limitations in the empirical literature and identify two types of patterns—the overall trajectory shapes and school stage transitions—driven by the two aspects of rural-urban divide.

GMM focuses on the shape of individuals' educational progress and transitions as a whole and identifies a finite number of distinct trajectory classes, technically "latent" classes (latent for unobserved). Two analytic patterns can be obtained. First, estimates for the most-likely membership of individuals in a latent class give the overall distribution of educational trajectories, which take into account not only the ultimate formal educational attainment but also how individuals' transitions pass through educational stages. Second, the properties of the distinct trajectory shape (e.g., the slope of growth in each educational stage) can be used to understand the likelihood of stage transitioning. For a formal introduction to GMM and its application to the current study, see the Appendix.

The event-history-calendar (EHC) information from CGSS 2008 provides our data. The EHC asked respondents to recall the beginning and ending years of each educational stage (primary, junior-high, senior-high or vocational, postsecondary vocational, 4-year college, and postgraduate), the place of the school (village, township, county seat, small city, prefectural city, provincial capital, and municipality), and the graduation status at each educational stage. Since educational experiences are salient to respondents, a benefit of the EHC calendar format is that it improves respondents' recall of the timing of transitions, helping them see the sequence of education in their lives.

From this data. we constructed our dependent variables, a set of 17 measures of years of schooling at each age from 6 to 22, which form the trajectory of years of schooling. We selected sample individuals who had passed their 22rd birthday and were not older than 64 in 2008, who were born between 1944 and 1986, and who had experienced education in socialist China. We truncated educational history at age 22, an age when most people with tertiary education have completed vocational or 4-year college education. If a person obtained their highest level of schooling before age 22, we extend this level to age 22. For example, for a person whose highest educational attainment was graduating from senior-high school at age 18, his trajectory will show an upward slope from 0 years of schooling at age 6 and 12 years of schooling at age 18, and then a flat line at 18 years of schooling from age 18–22.

In constructing schooling trajectories, we standardize years of schooling as 6 for completing primary, 9 for completing junior high, 12 for completing senior high or secondary vocational, 14 for completing postsecondary technical, and 16 for completing 4-year

college, regardless of the actual duration. For example, if a person dropped out of junior high for a year and then returned to complete it in a 4th year, we assign an 0.75 year increment per elapsed year over her 4 years for junior high. Likewise, as some rural primary schools operate on a 5-year basis, we assign a 1.2 year increment per elapsed year as long as the person reported graduating from primary school.

A first key explanatory variable is at-birth *hukou* status. The CGSS asks the current *hukou* status and place, distinguishing between (1) agricultural (rural) *hukou* in villages (43.1%), (2) blue-stamp *hukou* in townships (5.8%), and (3) nonagricultural (urban) *hukou* (51%) in various administrative urban areas. Because blue-stamp *hukou* is a variant of rural *hukou*, given that it does not grant the right to purchase commercial grain, we include blue-stamp *hukou* in the rural *hukou* category. Thus, *hukou* is a dichotomous variable distinguishing between rural and urban. If a person changed his or her rural *hukou* to urban *hukou*, the survey collected the year of this change. This information allows us to determine each individual's at-birth *hukou*.

Converting one's rural *hukou* to urban *hukou* during childhood enhances the child's opportunities, which may or may not compensate for the at-birth rural *hukou* disadvantage. Because rural-to-urban *hukou* conversion has been given on the basis of college educational attainment and military officer rank attainment, both of which occur during adulthood, childhood *hukou* conversion occurs when the mother's rural *hukou* is converted or when villages and townships are enclosed in a city's boundary, causing a blanket *hukou* conversion of the involved rural population. Using the information on the timing when a respondent's *hukou* was converted, we constructed an indicator for childhood *hukou* conversion – rural-to-urban *hukou* change by age 18 – to capture the change in opportunities.[1]

Another key explanatory variable is pre-tertiary school location, a dichotomous variable indicating rural vs. urban pre-tertiary school. The location of the attended school in the EHC lists village, township, county seat, and various administrative levels of cities. While the county department of education is the lowest educational administration governing rural education, some county schools are funded by prefectural or provincial governments. Thus, we consider schools in villages and townships as rural, and schools in county seats and higher administrative places as urban. For respondents whose educational attainment is lower than senior high, we consider whether their attended schools were all rural up to their educational attainment level.

Although Chinese educational stratification research commonly examines the effects of family origins on educational inequality, the theoretical rationale typically adopted is more relevant to developed societies, where social status groups defined by parental education, occupation, and income (SES) are stable over time. However, family SES in China does not capture the same stable social status groups as those in developed societies. For instance, Chinese parents with a Ph.D. were defined as being in a low social position by the class lines up to 1979. Yet, family SES still plays a role in family processes. Parents with higher

[1]There is virtually no urban-to-rural hukou change.

education and occupation may have higher educational expectations for their children, invest more in their children's education, and encourage their children to attain higher education. In addition, family structure, measured by parental widowhood by the time the respondent reached age 14, does not have the same meaning as single-parent families in developed societies; instead, it captures the origin family's poverty status. Parental party membership indicates a higher social-power position. Thus, we use family-origin variables to control for other factors of educational trajectories. We also control for gender, another main stratification factor, and birth order, which represents the cultural value on the first born or the only child.

To capture the broad educational development and historical contexts, we divided the sample into three cohorts who were born in 1944–1960, 1961–1970, and 1971–1986. The educational expansions in the past half-century exposed these cohorts to different opportunities. For example, the oldest cohort was too old to enjoy the educational expansion to junior high when they reached age 12 during 1956–72; in contrast, the youngest cohort benefited from the junior-high compulsory education at age 15 in 1986–2001, under the compulsory education law of 1986. Thus, the cohort variable captures the different educational opportunities as individuals were passing through their educational stages. Also controlled is regional-development imbalance, captured by an individual's current residence in the Eastern, Central, or Western regions, which is also an indicator for those who might have lived in a different region before age 14.

We intended to characterize the shape of educational trajectories corresponding to the typical transitioning from primary to junior-high, to senior-high, and then to tertiary. In statistical terms, this translates into a three-knot linear spline with the knots at age 12, 15, and 18. We have tested this time function and found it to be optimal among a set of time functions with different numbers and positions of knots and a set of polynomial time functions.

GMM is used to identify distinct latent classes of trajectories. Figure 1 is the graphical representation of the model, where $Y_0, Y_1, Y_2, \ldots, Y_{16}$ stand for the observed dependent variables (the years of schooling at each of the 17 time points from age 6–22, a non-decreasing function); I and $S_1, \ldots, S_4$ stand for the intercept and the linear spline slopes (growth factors); C is a categorical variable for a finite number of latent classes of trajectories; and X is a vector of explanatory variables, including at-birth *hukou*, rural-to-urban *hukou* change by 18, location of pre-tertiary school, family origin, individual characteristics, birth cohort, and region. The diagram shows that the observed dependent variables are characterized by growth factors within latent classes into which individuals are sorted by the vector of covariates. The last two slopes with latent classes are also a function of X. The model allows for unobserved individual heterogeneity, shown as an arrow to I, the intercepts.

To select the most parsimonious and good-fitting model, we used both statistical and substantive criteria. Two statistical criteria included the smallest value of Bayesian Information Criterion (BIC) and a significant test of the Lo-Mendell-Rubin (LMR) adjusted likelihood ratio test that compares the current model and a model with one fewer latent

class. We fitted a series of models from two latent classes to five latent classes. Our estimates for the five-class model were unstable. The four-class model showed a great improvement of goodness-of-fit and substantively meaningful distinct latent classes of trajectories—the highest, next-to-highest, next-lowest, and lowest. Therefore, our choice of the final was the four-class model.

Results

To situate our latent trajectory analysis in observed patterns, we first examined educational attainment by the two aspects of rural-urban divide. Table 1 shows sharp disparities between rural and urban *hukou*, as well as between exposure to rural and urban pre-tertiary schools. Individuals with no formal schooling account for 10.8% of the rural *hukou* group and 1.0% of the urban *hukou* group. The share of those attaining only primary schooling is also much greater for rural *hukou* (31.9%) than for urban *hukou* (7.8%). Taken at-or-below junior-high schooling together, the share (74.6%) of those with rural *hukou* is double the share (37.3%) of their urban-*hukou* counterparts. At the two above-junior-high levels, the share for rural *hukou* is much smaller than that for urban *hukou*, and the total above-compulsory share is 25.4% for rural *hukou* and 62.7% for urban *hukou*. Although the last two columns show that the disparities by exposure to rural-urban pre-tertiary schools are similar to the disparities by *hukou*, the gaps between rural and urban schools are larger: the at-or-below junior-high gap is between 83.8% (rural) and 29.5% (urban) and the above junior-high gap is between 16.2% (rural) and 70.5% (urban). This difference in the relationship between the two aspects of rural-urban divide makes it possible for us to tease out their respective stratifying effects in a multivariate manner.

To view these patterns visually, we present Figure 2, which shows the percentage distribution of the cross-classification of the two aspects of the rural-urban divide in the seven levels of educational attainment. The middle two segments are for the share of the two mixed categories. The patterns can be used to address two issues: (1) the degree to which the two aspects of rural-urban divide are not overlapping, and (2) the degree to which the two aspects of rural-urban divide are associated with educational attainment.

Regarding the degree of non-overlap between the two aspects of the rural-urban divide, we focus on the middle two segments. The share for the U-R mixed cases (urban *hukou* and rural pre-tertiary school) is smaller than the share for the R-U mixed cases (rural *hukou* and urban pre-tertiary school). For instance, about three to five percent of urban-*hukou* people attended rural primary and junior-high schools, and this percentage is relatively stable throughout levels of education. In comparison, the percentage for those with rural *hukou* who attended urban pre-tertiary schools is very small at primary and junior-high educational attainment but increases at the senior-high and above levels of attainment (more than 20%). These substantial proportions of mixed cases facilitate our estimation of the separate effects of rural-urban *hukou* and rural-urban pre-tertiary schools.

Regarding the association between the rural-urban divide and educational attainment, we see that the proportion of the U-U cases (urban *hukou* and urban school) increases with attainment levels but the proportion of R-R cases (rural *hukou* and rural school) declines

with attainment levels. In addition, the proportion of R-U increases with attainment levels, but the proportion of U-R mixed cases is stably low. These patterns suggest differential roles of the two aspects of rural-urban divide, which will be examined in the GMM analysis.

Returning to Table 1, we observe that people with rural-*hukou* are similar to people attending rural pre-tertiary schools in their disadvantages of family origins: lower levels of parental education, lower percentages of parents with professional occupations and parental party membership, and higher numbers for widowhood. Both the percentage of males and the percentage of first-born or only child are higher among urban-*hukou* people or among people attending urban pre-tertiary school. The cohort percentage distribution, like the region percentage distribution, differs little by rural-urban *hukou* or by rural-urban school.

Turning to the trajectory analysis results, we first examine the shape and distribution of the four latent classes estimated from the GMM analysis. In Figure 3, the x-axis indicates time points from 0 to 16, corresponding to ages 6–22; the y-axis indicates years of schooling. The three knots of the linear splining function are at time points 6, 9, and 12 (corresponding to ages 12, 15, and 18).[2] The figure shows the mean estimated intercept and slopes for each latent class of educational trajectories. As years of schooling are cumulative in nature, the trajectories are non-decreasing. The latent class for the highest trajectory includes a majority of those who reached senior-high or tertiary education, accounting for 40.8% of the sample. The latent class for the next-to-highest trajectory characterizes mainly those who completed compulsory education but not higher, accounting for 28.9% of the sample. The latent class for the next-to-lowest trajectory represents the majority of those who finished only primary education, accounting for 17.4%. Finally, the latent class for the lowest trajectory captures mainly those with no formal education and primary school dropouts, accounting for 13.0% of the sample. The four distinct mean trajectories diverge over the ages 6–22.

Emerging from the observed years of schooling at each of the 17 time points from age 6 to 22, and determined by the two aspects of rural-urban divide and other covariates, each individual has a probability of falling in one of the four latent classes. The estimated most-likely class membership is then determined for each respondent. The uncertainty of membership is quite small, as the mean most-likely probabilities range from 0.996 to 0.999. In the top panel of Table 2, we see how the latent classes do not completely overlap with the educational attainment levels, as all latent classes include at least three out of four collapsed levels of attainment.[3] For instance, 81.4% of the highest latent class attained senior high or above, and 17.7% completed junior high. Greater diversity is found among lower latent classes. The next-to-highest latent class includes 72.7% junior high and 15.8% senior high or above; the next-to-lowest latent class includes 84.8% primary and 13.7% junior high; and the lowest latent class includes 57.4% no formal education and 40.9 % primary. The latent classes more realistically reflect the number of years in and the completion of each stage, whereas the reported educational attainment does not discriminate between completion and dropout in a stage.

[2]The first linear spline stretches down to −1, which is a statistical artifact.
[3]This is because the timing and duration of completing each educational stage as well as other unobserved characteristics vary across individuals, all of which are taken into account in the GMM analysis.

We next examine the distribution of covariates across the latent classes of trajectories. The next three rows of Table 2 show increasing percentages of at-birth rural *hukou* and rural pre-tertiary schools from the highest to the lowest latent classes, indicating the strong stratification role of the two aspects of rural-urban divide. In contrast, the percentage of rural-to-urban *hukou* changes by age 18 declines from the highest to the lowest trajectory. A closer look, however, reveals that rural-urban schools discriminate among the four latent classes more than at-birth rural-urban *hukou*. In particular, the percentage of rural-*hukou* individuals in the four latent classes increases gradually from 51.4% in the highest, 62.4% in the next-to-highest, 85.6% in the next-to-lowest, to 93.3% in the lowest. The percentage of rural school is 38.5% in the highest, 57.4% in the next-to-highest, 87.7% in the next-to-lowest, and 93.5% in the lowest. These patterns suggest differential stratification roles played by the two aspects of rural-urban divide.

Patterns of family background are consistent with the at-birth *hukou* patterns: parental education, occupation, widowhood, and party membership are distributed monotonically with the rank of the four latent classes. These results suggest that the family process through which parents advantaged in education, occupation, two-parent structure, and power derived from the party membership invest more in their children's education. The gender effect is also similar to the *hukou* effects. Birth order, however, discriminates relatively little across the latent classes.

Cohorts capture the impact of educational expansions over the past 50 years and exhibit three patterns. The proportion of the oldest cohort monotonically increases from the highest to the lowest latent classes. The opposite is true for the youngest cohort. The middle cohort has a curvilinear trend. These patterns are consistent with the educational-expansion history from the primary expansion in the 1950s, to compulsory education in the 1980s, and to the latest college education expansion since 1999. The Eastern vs. Central or Western regional differences in educational development are as expected: people living in the less-developed Central and Western regions are more likely to fall into lower latent classes.

The GMM coefficients reveal whether the two aspects of rural-urban divide sort people into the four distinct latent classes while controlling for other covariates, and they provide evidence to support or reject our hypotheses. We estimate the effects of *hukou* and school simultaneously in the "*hukou*-and-school model", the results of which are compared with those from the "*hukou* model" (with school eliminated) and the "school model" (with *hukou* eliminated). The results are presented in Table 3. Within the models, we contrast each of the three top latent classes against the lowest latent class. The estimates are the log odds of falling in any other latent class vs. the lowest latent class.

We examine the estimates in Table 3 with three questions centering on the two aspects vs. one aspect of rural-urban divide. First, how different are the effects of the two aspects on individuals' memberships in the latent classes of trajectories? Second, how do these two effects compare with the effects of other covariates in the *hukou*-and-school model?

Third, how might the one-aspect analysis misspecify the model? To answer the first question, we examine the top two rows in the column for the *hukou*-and-school model.

When estimated simultaneously with school, at-birth rural *hukou* has a significant, negative effect on the membership in each of the top three latent classes referenced to the lowest latent class. Additional tests (not shown in the table) confirm that there is no significant difference in the effect on the memberships between the highest class and the next-to-highest class, both of which are strongly unlikely for people with at-birth rural *hukou*. Overall, individuals with at-birth rural *hukou* are less likely to fall into the top three trajectories, all else being equal. However, changing from rural to urban *hukou* during childhood greatly increased the odds of following the higher trajectories. For example, the positive coefficient (2.237) overcompensates for the detrimental effect of at-birth rural *hukou* (−1.213) for the highest trajectory class. A similar overcompensation is found for the next-to-highest class. While the large-size effect of childhood *hukou* change for the next-to-lowest class is not significant, this may be due to the small cell size of respondents with childhood *hukou* change in the next-to-lowest class.

The effects of rural pre-tertiary school are also negative, but they are different from the effects of at-birth rural *hukou* in a notable way: rural schools significantly discriminate among the four latent classes. In other words, exposure to rural pre-tertiary school lowers the probability of being in the highest class, followed by the next-to-highest and then the next-to-lowest, and finally the lowest. Comparing the two aspects of rural-urban divide, the detriment of rural school for the top two trajectory classes is stronger, at −2.039 and −1.427, than the detriment of at-birth rural *hukou*, at −1.213 and −1.233. But it is at-birth rural *hukou*, rather than rural school, that is responsible for keeping rural-*hukou* people in the lowest latent class and thus at the bottom of educational stratification.

To answer the second question—the two effects compared with the effects of other covariates in the *hukou*-and-school model—we see that the two aspects of rural-urban divide are among the strongest factors in determining memberships in the higher latent classes as compared to the lowest latent class. After rural *hukou* and rural school, risk factors include lower parental education and occupation, parental widowhood, parental non-party status, being female, being in the earliest cohort, and living in the Western region, which are what we expected.

Regarding the third question about one aspect vs. two aspects for model specification, we compare coefficients in the first three rows across models. Under the *hukou* model, the at-birth rural *hukou* coefficients are significant and larger than those in the *hukou*-and-school model. The rural pre-tertiary school coefficients from the school model show a similar pattern. These results suggest that the two aspects of rural-urban divide confound with one another to some degree but they do not completely overlap. Therefore, leaving out either one leads to model misspecification and distorts the role of rural-urban divide in reproducing educational inequality.

Taking together our answers to the three questions, we examine evidence for our hypothesis testing. The *hukou*-school model provides strong evidence to support our first hypothesis that at-birth *hukou* sorts rural-*hukou* people to the lowest trajectory class. The significant negative effects of at-birth rural *hukou* on all three higher trajectory classes when contrasted with the lowest class suggest that this ascribed social status is indeed responsible for

defining the bottom of educational stratification. At the same time, the *hukou*-school model shows strong support for the second hypothesis that rural pre-tertiary school penalizes the students' chance to get ahead, especially to follow the highest trajectory class compared to the lower three latent classes. It is the *hukou*-school model that teases out the differential effects between *hukou* and school. The *hukou* model tells nothing about the school effect. The school model tells nothing about the *hukou* effect and also makes the school effect appear responsible for keeping rural school students in the very bottom of the educational stratification.

Hypothesis 3 is about the cumulative disadvantage of *hukou* and school and the hypothesized divergent trajectories. Figure 3, based on the *hukou*-school model, shows that the four distinct mean trajectories increasingly diverge with educational stages. Because rural-*hukou* people are more likely in the lowest class and rural-school students in the lowest and next-to-lowest classes, whereas urban-*hukou* people and urban-school students are more likely in the highest and next-to-highest classes, the evidence is strong to support Hypothesis 3 that the disadvantage of rural pre-tertiary school is greater for later educational stages than earlier ones.

Additional tests for Hypothesis 3 are from the differential slopes of the senior-high and tertiary stages within each latent trajectory class as a function of the same vector of covariates as that for the latent trajectory classes. The effects of covariates on the slopes are above and beyond the effects of the same covariates on the latent classes. Table 4 reports the estimates for the two aspects of rural-urban divide on the two sets of slopes. The estimates show that rural *hukou* has a stronger negative effect on the growth over the senior-high stage within the next-to-lowest class than the highest class. These effects are rather sporadic. More systematic findings are the negative effects of rural pre-tertiary schools: stronger on the slope over the tertiary stage than the senior-high stage within the highest class, where more individuals transitioned to the tertiary stage; and stronger on the slope over the senior-high stage than the tertiary stage with the next-to-highest class, where few individuals transitioned to the tertiary stage. Not surprisingly, no school effects are found for the two lower classes. The systematic negative school effects on the growth of years of schooling over the last two educational stages within the highest and next-to-highest classes further support our Hypothesis 3 about the cumulative disadvantage of exposure to rural pre-tertiary school.

Conclusions

Contextualized in China's social change in the past half-century, this paper develops the notion of dichotomous inequality, akin to what Tilly (1998) called categorical inequality, to conceptualize China's rural-urban divide in educational opportunities and outcomes. It argues that the *hukou* system assigns people to a top-bottom hierarchy and the rural-urban school system institutionalizes unequal resource distribution and diverse mission. The initial disadvantages of rural *hukou* and rural school accumulate over an individual's life course and contribute to divergent educational trajectories. Together, the *hukou* and school aspects of rural-urban divide have maintained educational inequality despite educational expansions in the past decades. With this conceptualization, we formulate a Chinese version of the MMI

in a dynamic perspective, under which we hypothesize that rural *hukou* holds its assignees at the very bottom of the educational stratification; the rural-urban educational system transfers unequal opportunities from lower educational stages to higher ones and determines who is more likely to follow the highest educational trajectory; and the disadvantage of rural *hukou* and rural school accumulate over time.

This conceptualization not only contributes to a better understanding of educational stratification in China, where the rural-urban development gap is enormous and similar to other developing societies with geographically imbalanced development, but also sheds new light on the advancement of educational stratification theories originated in the developed world. While scholars usually agree on the detrimental structural effects of social class and poverty, consensus is yet to be achieved about school effects. Explanations of school effects compete between the reproduction of inequality argument (Bowels and Gintis 1976), the summer learning loss explanation (Alexander et al. 2001), and the effectively maintained inequality-by-tracking theory (Lucas 2001). Past research has usually conceived school effects on a particular level of educational process (DiPrete and Eirich 2006). Our approach is to define a pre-tertiary school dichotomy without the loss of its institutionalized unequal resource allocation and diverse mission. Our trajectory view pays attention to school effects on the whole trajectory and various educational stages over the trajectory. The differential consequences of exposure to the negative treatment of rural pre-tertiary school can be evaluated with respect to distinct types of trajectories as well as successive educational stages. Our paper suggests that defining the school institution as a whole and evaluating this school institutional effect on the whole trajectory as well as on successive educational stages may be effective to establish a consensus on school effects.

Our paper is methodologically innovative in implementing its conceptual ideas. Capitalizing on the available individual educational history calendars, we construct individual educational trajectories of years of schooling from the cross-sectional survey of CGSS 2008. We estimate generalized mixture models to provide analysis for the overall stratification of educational trajectories and the educational stage transitions, both of which has been addressed separately in the empirical literature. Rather than analyzing cross-sectional educational attainment or the transition from one stage to the next, our trajectory analysis allows both a holistic view of the trajectory shape and a microscopic view at each time point of a trajectory.

With these theoretical and methodological improvements, we offer three major findings. First, we find strong evidence in support of *hukou*'s sorting mechanism. Individuals with rural *hukou* assigned at birth are more likely to follow the lowest educational trajectory than their urban-*hukou* counterparts, which could be offset by childhood *hukou* change, a small-probability event. Second, the evidence is strong and systematic that exposure to rural pre-tertiary school penalizes the transition to a higher stage, and this penalty increases significantly with levels of educational stages and becomes the strongest for transitioning to the tertiary stage. Because the privileged urban-*hukou* people have saturated the lower stages, the opportunity hoarding accelerates with educational stages, enabled by the school institutionalization of unequal resources allocation and diverse mission. Third, exposure to rural *hukou* and rural pre-tertiary school compounds the disadvantage, exhibited in the

divergence of our identified four distinct educational trajectories. The detrimental effects of rural school, rather than rural *hukou*, are significantly stronger for later educational stages. These findings taken together attest to the Chinese-version MMI and inequality reproduction even during an educational expansion. Overall, these findings highlight the importance of both the structural factor of *hukou* and the institutional factor of school, the two aspects of rural-urban divide, in reproducing educational inequality in China. While both effects are detrimental, they are notably different with different mechanisms. Overlooking either aspect of rural-urban divide will obscure our understanding of the sources and mechanisms of educational inequality.

Our findings about how educational inequality is reproduced can inform policymakers intending to increase social mobility. The fact that both aspects of the rural-urban divide play a vital but differential role in such a reproduction suggests three policy options. First, abolishing the *hukou* system is a straightforward one, because all at once it would eliminate the hierarchy it previously created and renewed generation after generation. The call to abolish the *hukou* system has been raised (Chan 2009), and our evidence reinforces the urgent need for solving the problem at its root. If the institutional change of the rural-urban educational system lacks or lags, however, inequality will still be continuously reproduced. This leads to our second policy option: to reform the rural-urban educational system and close the gap in resource allocation and school mission. Setting the same school mission and increasing the central government's resource allocation to rural schools may be two measures. Yet, if *hukou* continues, it will constantly create pressure to institutionalize unequal resource allocation, undermining reform efforts. The limitations of the first two policy options suggest the best option: a policy that combines the first and second options and removes both aspects of the rural-urban divide.

References

Alexander, Karl L.; Entwisle, Doris R.; Olson, Linda S. Schools, achievement, and inequality: A seasonal perspective. Educational Evaluation and Policy Analysis. 2001; 23:171–91.

Blau PM, Duncan OD. The American occupational structure. 1967

Boli J, Ramirez FO, Meyer JW. Explaining the origins and expansion of mass education. Comparative Education Review. 1985; 29:145–170.

Bowels, S.; Gintis, H. Schooling in capitalist America: Educational reform and the contradictions of economic life. New York: Basic Books; 1976.

Breen R, Goldthorpe JH. Explaining educational differentials towards a formal rational action theory. Rationality and society. 1997; 9:275–305.

Breen R, Jonsson JO. Analyzing educational careers: A multinomial transition model. American Sociological Review. 2000:754–772.

Breen R, Jonsson JO. Inequality of opportunity in comparative perspective: Recent research on educational attainment and social mobility. Annual review of sociology. 2005:223–243.

Cameron, SV.; Heckman, JJ. Life cycle schooling and dynamic selection bias: Models and evidence for five cohorts. National Bureau of Economic Research; 1998.

Chan, KW. Cities with invisible walls: reinterpreting urbanization in post-1949 China. Oxford University Press; Hong Kong: 1994.

Chan KW. The Chinese hukou system at 50. Eurasian Geography and Economics. 2009; 50:197–221.

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

DiPrete TA, Eirich GM. Cumulative advantage as a mechanism for inequality: A review of theoretical and empirical developments. Annual Review of Sociology. 2006:271–97.

Duncan GJ, Yeung WJ, Brooks-Gunn WJJ, Smith JR. How much does childhood poverty affect the life chances of children? American Sociological Review. 1998; 63:406–23.

Elder, GH. Life course dynamics: Trajectories and transitions, 1968–1980. Cornell University Press; Ithaca: 1985.

Erikson R, Jonsson JO. Explaining class inequality in education: The Swedish test case. Can education be equalized. 1996:1–63.

Fu TM. Unequal primary education opportunities in rural and urban China. China perpectives. 2005

Guo C, Tsang MC, Ding X. Gender disparities in science and engineering in Chinese universities. Economics of Education Review. 2010; 29:225–235.

Hannum E. Political change and the urban-rural gap in basic education in China, 1949–1990. Comparative Education Review. 1999; 43:193–211.

Hannum E, Park A, Cheng K. Introduction: Market reforms and educational opportunity in China. Education and reform in China. 2007:1–23.

Hannum E, Wang M. Geography and educational inequality in China. China Economic Review. 2006; 17:253–265.

Hannum, E.; Xie, Y. Trends in educational gender inequality in China: 1949–1985. University of Michigan; 1994.

Hauser RM, Featherman DL. Equality of schooling: Trends and prospects. Sociology of Education. 1976:99–120.

Knight J, Shi L. Educational Attainment and the Rural Urban Divide in China. Oxford Bulletin of Economics and Statistics. 1996; 58:83–117.

Lee, MDP. Widening Gap of Educational Opportunity. World Institute for Development Economics Research, United Nations University; 2006.

Li, Chunling. The Expansion of Higher Education and the Inequality of Educational Opportunity: An Empirical Analysis of Equalization Effect (Gaodeng Jiaoyu Kuozhang yu Jiaoyu Jihui Bupingdeng: Gaoxiao Kuozhao de Pingdenghua Xiaoying Kaocha). Shehuixue Yanjiu. 2010; 3:83–113.

Lipset SM. Some social requisites of democracy: Economic development and political legitimacy. The American Political Science Review. 1959:69–105.

Lucas, SR. Tracking Inequality: Stratification and Mobility in American High Schools. New York: Teachers' College; 1999.

Lucas SR. Effectively Maintained Inequality: Education Transitions, Track Mobility, and Social Background Effects1. American Journal of Sociology. 2001; 106:1642–1690.

Mare RD. Social background and school continuation decisions. Journal of the American Statistical Association. 1980; 75:295–305.

Mare RD. Change and stability in educational stratification. American Sociological Review. 1981:72–87.

Massey, DS. Categorically unequal: The American stratification system. Russell Sage Foundation Publications; 2007.

McLanahan, S.; Sandefur, G. Growing Up with a Single Parent: What Hurts, What Helps. Cambridge, MA: Harvard University Press; 1994.

McLeod JD, Fettes DL. Trajectories of failure: The educational careers of children with mental health problems. AJS; American journal of sociology. 2007; 113:653.

Ministry of Education of China. 2012. http://www.moe.edu.cn/publicfiles/business/htmlfiles/moe/moe_2792/index.html

Morgan SL. Adolescent educational expecations rationalized, fantasized, or both? Rationality and society. 1998; 10:131–162.

Muthén, B. The Sage handbook of quantitative methodology for the social sciences. Thousand Oaks, CA: Sage Publications; 2004. Latent variable analysis; p. 345-68.

Muthén B, Shedden K. Finite mixture modeling with mixture outcomes using the EM algorithm. Biometrics. 1999; 55:463–469. [PubMed: 11318201]

Pepper S. Chinese education after Mao: two steps forward, two steps back and begin again? The China Quarterly. 1980; 81:1–65.

Raftery AE, Hout M. Maximally maintained inequality: Expansion, reform, and opportunity in Irish education, 1921–75. Sociology of Education. 1993:41–62.

Shavit, Yossi; Blossfeld, Hans-Peter. Persistent Inequality: Changing Educational Attainment in Thirteen Countries. Social Inequality Series. ERIC; 1993. Persisting Barriers: Changes in Educational Opportunities in Thirteen Countries.

Stoolmiller M, Kim HK, Capaldi DM. The course of depressive symptoms in men from early adolescence to young adulthood: identifying latent trajectories and early predictors. Journal of Abnormal Psychology; Journal of Abnormal Psychology. 2005; 114:331.

Szelenyi, I. Urban inequalities under state socialism. Oxford University Press; Oxford: 1983.

Tilly JL. Molecular and genetic basis of normal and toxicant-induced apoptosis in female germ cells. Toxicology letters. 1998; 102:497–501. [PubMed: 10022302]

Titma M, Tuma NB. Stratification Research in a Changing World. Eastern European Societies on the Threshold of Change. 1993:225–254.

Tsang MC. Education and national development in China since 1949: Oscillating policies and enduring dilemmas. China review. 2000; 2000:579–618.

Whyte, MK.; Parish, WL. Urban life in contemporary China. University of Chicago Press; 1985.

Wu X. Economic transition, school expansion and educational inequality in china, 1990–2000. Research in Social Stratification and Mobility. 2010; 28:91–108.

Wu X, Treiman DJ. The household registration system and social stratification in China: 1955–1996. Demography. 2004; 41:363–384. [PubMed: 15209045]

Wu X, Treiman DJ. Inequality and Equality under Chinese Socialism: The Hukou System and Intergenerational Occupational Mobility1. American Journal of Sociology. 2007; 113:415–445.

Zhang X, Kanbur R. Spatial inequality in education and health care in China. China Economic Review. 2005; 16:189–204.

Zhou X, Moen P, Tuma NB. Educational stratification in urban China: 1949–94. Sociology of Education. 1998:199–222.

Appendix. Generalized Mixture Modeling Applied to Educational Trajectories

This paper applies the generalized mixture modeling (GMM) to individual trajectories of years of schooling over ages 6–22 among 5,187 sample individuals from the China General Social Survey (CGSS) 2008. This appendix gives a brief formal introduction to GMM. Let y_{it} be the repeated observations of years of schooling for individual i at time t, t = 0,1,…,16 for ages 6–22. Given different educational stages and transitions, we use a three-knot linear spline growth function with the knots at *age* = 6,9,12. Let X_i be a vector of covariates. We propose that the distribution of individual educational trajectories is a mixture of K distinct latent classes. Let c_{ki} be a set of indicators (dummy variables) with value 1 for individual i's membership in the k th class (k = 1,…, K −1) and 0 otherwise. Our GMM is expressed in equations (1)–(3):

$$y_{it}=\sum_{k=1}^{K-1}(c_{ki}=1)[\beta_{0ki}+\beta_{1ki}t_1+\beta_{2ki}t_2+\beta_{3ki}t_3+\beta_{4ki}t_4+\varepsilon_{ikt}] \quad (1)$$

$$\begin{aligned} \beta_{0ki} &= \gamma_{00k} + u_{0ki} \\ \beta_{qki} &= \gamma_{q0k} \quad q = 1, 2 \\ \beta_{qki} &= \gamma_{q0k} + \gamma_{q1k} X_i \quad q = 3, 4 \end{aligned} \tag{2}$$

$$\log\left(\frac{p(c_i = k)}{p(c_i = K)}\right) = \delta_{0k} + \delta_{1k} X_i \tag{3}$$

Equation (1) states that the individual trajectories form a mixture of *K* latent classes, within each of which a three-knot linear-spline time function (four slopes) is specified. Equation (2) shows that the growth factors (intercepts and slopes) differ across the *K* distributions. Equation (2) allows the 3rd and 4th sets of slopes to condition on X_i and specifies individual heterogeneity in the intercepts (u_{0ki}). Equation (3) assumes that the probability for individual *i* to fall into the *k* th latent class is conditioned on X_i. The estimation of (1)–(3) uses the EM algorithm implemented in Mplus 7.0.

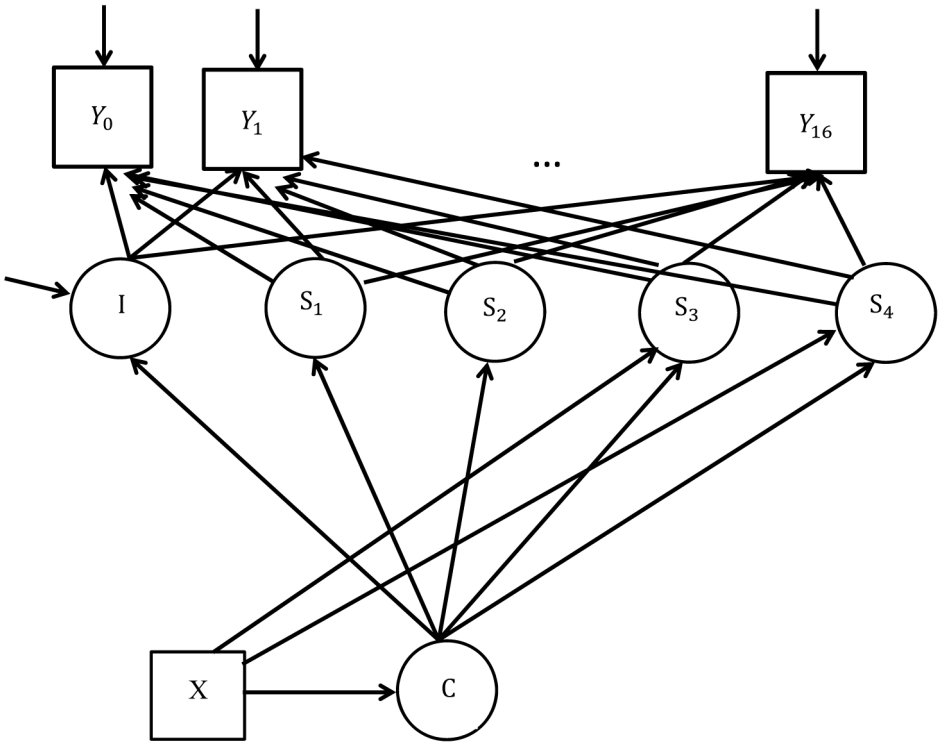

Figure 1.

A Generalized Mixture Model Application to Educational Trajectories

Note: Y's are years of schooling at each age from 6 to 22, I is the intercepts, S's are linear spline slopes, C is a categorical variable indicating the latent classes of trajectories, and X is a vector of covariates.

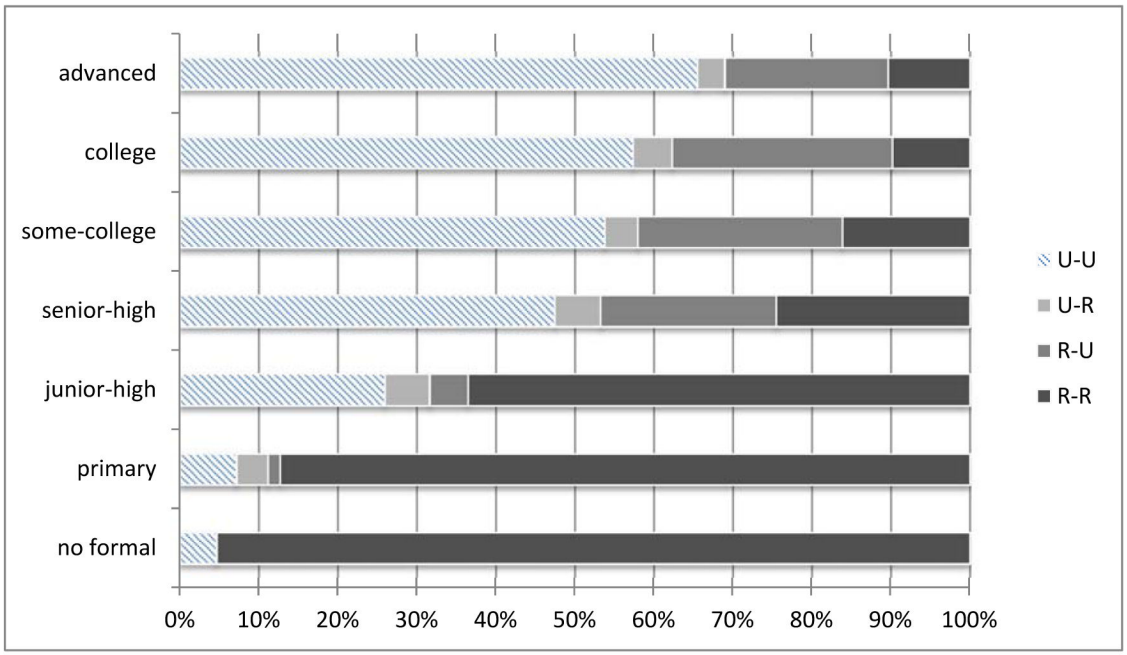

Figure 2.

Percentage Distribution of At-birth *Hukou* and Pre-Tertiary School within Educational Attainment Levels

Note:

U-U: at-birth urban *hukou* & urban pre-tertiary school

U-R: at-birth urban *hukou* & rural pre-tertiary school

R-U: at-birth rural *hukou* & urban pre-tertiary school

R-R: at-birth rural *hukou* & rural pre-tertiary school

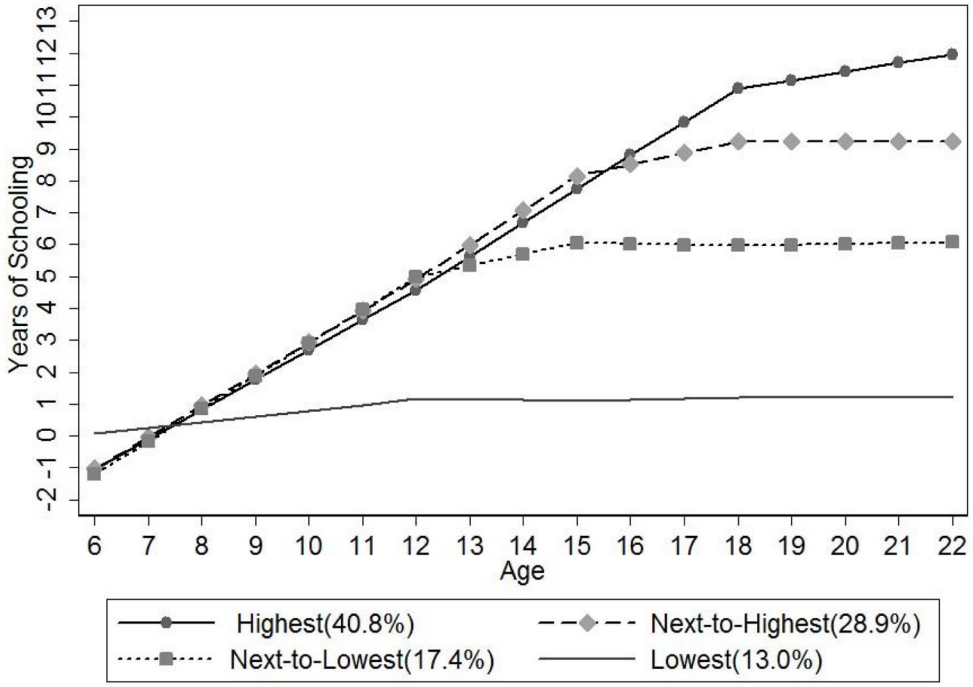

Figure 3.
Four Latent Classes of Educational Trajectories: Mean Estimated Intercepts and Slopes based on the Two-Aspect Model

Table 1

Observed Educational Attainment by Two Aspects of Rural-Urban Divide

Variable	At-birth *Hukou*		Pre-tertiary School	
	Urban	Rural	Urban	Rural
Educational attainment				
No formal schooling	0.010	0.108	0.009	0.119
Primary school	0.078	0.319	0.051	0.363
Junior high school	0.285	0.319	0.235	0.356
Senior high school	0.360	0.163	0.398	0.117
Tertiary	0.267	0.091	0.308	0.045
Family origin				
Parental education (years)	7.145	4.297	7.295	3.894
Parental professional occupation	0.376	0.099	0.353	0.086
Parent died by 14	0.091	0.107	0.075	0.119
Parent party membership	0.195	0.112	0.198	0.101
Individual characteristics				
Male	0.500	0.469	0.508	0.460
First or only child	0.401	0.292	0.400	0.281
Cohort				
Born 1944–1960	0.355	0.329	0.302	0.363
Born 1961–1970	0.265	0.300	0.258	0.309
Born 1971–1986	0.380	0.370	0.441	0.328
Region				
Eastern	0.449	0.295	0.410	0.305
Central	0.356	0.445	0.372	0.444
Western	0.195	0.260	0.218	0.251
Lived in a different region by 14	0.112	0.283	0.204	0.238
n	1766	3421	2093	3094

Source: Authors' compilation using CGSS 2008.

Table 2

Characteristics of Members in Latent Classes of Educational Trajectories: Two-Aspect Model

Characteristic	Latent Class			
	Highest	Next-to-Highest	Next-to-Lowest	Lowest
Observed educational attainment				
No formal schooling	0.000	0.000	0.000	0.574
Primary school	0.009	0.114	0.848	0.409
Junior high school	0.177	0.727	0.137	0.010
Senior high school and above	0.814	0.158	0.014	0.006
Rural-urban divide				
At-birth rural *hukou*	0.514	0.624	0.856	0.933
Rural-to-urban *hukou* change by 18	0.049	0.030	0.014	0.004
Rural pre-tertiary school	0.385	0.574	0.877	0.935
Family origin				
Parental education (years)	7.053	5.336	3.504	1.854
Parental professional occupation	0.293	0.193	0.083	0.028
Parent died by 14	0.068	0.099	0.114	0.193
Parent party membership	0.193	0.134	0.091	0.054
Individual characteristics				
Male	0.538	0.487	0.468	0.298
First or only child	0.379	0.308	0.283	0.280
Cohort				
Born 1944–60	0.260	0.300	0.380	0.612
Born 1961–70	0.251	0.344	0.330	0.225
Born 1971–86	0.488	0.356	0.289	0.164
Region				
Eastern	0.439	0.321	0.267	0.226
Central	0.362	0.441	0.490	0.423
Western	0.199	0.239	0.243	0.351
Lived in a different region by 14	0.255	0.199	0.200	0.217
n	2117	1496	902	672

Note: Estimated latent class memberships from the two-aspect model in Table 3.

Table 3

Estimates of Latent Trajectory Classes: Two-Aspect and one-Aspect Models (the lowest trajectory class as the reference)

Variable	*Hukou*-and-School Model			*Hukou* Model			School Model		
	highest	next-to-highest	next-to-lowest	highest	next-to-highest	next-to-lowest	highest	next-to-highest	next-to-lowest
Rural-urban divide									
At-birth rural *hukou*	−1.213***	−1.233***	−0.858***	−2.712***	−2.183***	−0.988***	--	--	--
Rural-to-urban *hukou* change by 18	2.237***	1.980**	1.203	2.752***	2.121**	1.193	--	--	--
Rural pre-tertiary school	−2.039***	−1.427***	−0.134	--	--	--	−2.806***	−2.213***	−0.687 ***
Family origin									
Parental education (years)	0.169***	0.126***	0.071***	0.182***	0.136***	0.071***	0.171***	0.128***	0.072***
Parental professional occ.	1.111***	0.977***	0.724**	1.368***	1.144***	0.856**	1.196***	1.058***	0.762**
Parent died by 14	−0.565**	−0.388**	−0.402**	−0.512**	−0.467**	−0.382*	−0.528**	−0.351*	−0.377 *
Parent party membership	0.503*	0.251	0.179	0.523*	0.274	0.203	0.515*	0.261	0.194
Individual characteristics									
Male	1.436***	1.170***	0.968***	1.402***	1.272***	0.983***	1.427***	1.157***	0.949***
First or only child	−0.036	−0.130	−0.038	0.046	−0.131	−0.039	−0.015	−0.111	−0.021
Cohort (born 1961–70 is reference)									
Born 1944–60	−1.022***	−1.208***	−0.907***	−1.159***	−1.099***	−0.912***	−0.945***	−1.132***	−0.865 ***
Born 1971–86	0.517**	0.070	0.075	0.449**	0.200	0.073	0.532**	0.077	0.076
Region (Eastern is reference)									
Central	−0.563***	−0.124	0.057	−0.483***	−0.184	0.053	−0.626***	−0.199	0.020
Western	−1.072***	−0.597***	−0.487**	−0.880***	−0.683***	−0.498**	−1.155***	−0.689***	−0.532 ***
Lived in a different region by 14	0.770***	0.403**	0.223	0.780***	0.610***	0.222	0.717***	0.350**	0.185

Note: Estimates from generalized mixture models.

* p < 0.05

** p < 0.001

*** p<0.000-

Table 4

Effects of Two Aspects of Rural-Urban Divide on Slopes over Senior-High and Tertiary Stages

Variable	Latent Trajectory Classes			
	Highest	Next-to-highest	Next-to-lowest	Lowest
Slope over the senior-high stage				
At-birth rural *hukou*	−0.023 *	−0.027	−0.065 **	−0.028
Rural-to-urban *hukou* change by 18	0.025	0.061	0.128	0.394
Rural pre-tertiary school	−0.049 ***	−0.186 ***	−0.016	−0.012
Slope over the tertiary stage				
At-birth rural *hukou*	−0.020	0.029	0.033	0.030
Rural-to-urban *hukou* change by 18	0.024	0.029	0.028	0.290
Rural pre-tertiary school	−0.236 ***	−0.065 **	−0.029	−0.179

Note: Estimated from the two-aspect model in Table 3. The model controls for other variables (family origin, individual characteristics, birth cohort, and geographic area).

* $p < 0.05$

** $p < 0.001$

*** $p<0.0001$