Educational Trends, 1999 Higher Education Expansion,
and Relative Intergenerational Occupation Persistence in Postsocialist China

by

Haorui Peng

A thesis submitted to
Department of Sociology, University of Michigan
in partial fulfillment of the requirements for the degree of
Bachelor of Arts with Honors

April 2025

Thesis Advisor:

Professor Alexandra A. Killewald

Program Coordinator:

Professor Karin A. Martin
Professor Renee R. Anspach

Abstract

Children of parents with high-status occupations often end up working in high-status occupations, while their peers whose parents have low-status jobs often end up in low-status jobs. Education plays a pivotal and multifaceted role in this relative intergenerational occupation persistence. Rapid social and educational change in China, and especially the 1999 Higher Education Expansion, may have produced changes in education's role in intergenerational persistence. I hypothesize that stratified higher education and occupation outcomes in the population may have strengthened intergenerational persistence, while growing percentages of college graduates, for whom occupational attainment tends to vary less by social origin, may have weakened persistence. Analyzing data from nine waves of Chinese General Social Survey, I first confirm that intergenerational occupational persistence in China remained stable between the 1964-1979 and 1980-1995 birth cohorts. Next, model-based simulations suggest that changes in (1) the association between parents' occupation status and children's educational attainment and (2) the correlation between children's education and occupation status in adulthood decreased persistence, while changes in (3) the gaps in adult occupation status by parental occupation status among people with the same education and (4) the shares of the population in different educational groups increased persistence. Overall cross-cohort stability in intergenerational occupation persistence resulted from offsetting contributions of these four educational processes. However, the analyses are underpowered, and further efforts are needed to draw more definitive conclusions when new data are available.

Acknowledgements

"We social scientists of the current generation are fortunate in being able to witness firsthand social changes in China and to have opportunities to observe, document, analyze, and understand them as they occur."
– Yu Xie (2011:17)

I can never agree more on this comment. I am lucky, that I am wrapping up my undergraduate career by taking my first formal step forward on studying China and identifying my life-long passion; I am even more fortunate, that I got to know so many fantastic individuals via my thesis project.

My thanks are due first and foremost to my thesis advisor, Alexandra Killewald, whose guidance and patience made this thesis possible. I am deeply grateful to Sasha for her feedback and suggestions on literally everything, from topic choice to quantitative methods, from logic to language in writing. Without her millions of comments, I could not grow to who I am today.

Thanks are also due to our program coordinators, Karin Martin and Renee Anspach, for organizing the wonderful thesis seminars. I have particularly benefited from Karin's comments that pointed out the places where I needed more clarification.

I have also benefited from my peers in the 2024-2025 senior thesis cohort and in the Sociology 542 class, who made the lonely journey of academic studies a joyful and sweet period of time. I am thankful for their feedback that encouraged me to keep improving my thesis. I have been fortunate to receive support from Jingyang Meng and Yiran Xu both for and beyond the thesis project.

Numerous additional faculty and staff from the University of Michigan and other institutions provided helpful instructions, tips, and support in various stages of my thesis project. I wish to thank in particular Sarah Burgard, Thomas Drake, Judith Dyer, Lauren Eddelbuettel, Daniel Schneider, and Yu Xie.

The Chinese General Social Survey is primarily committed by the National Survey Research Center at the Renmin University of China (RUC), sponsored by RUC 985 Grant and RUC Scientific Research Grant.

Lastly, I thank all my friends and my family for their love and support. Special thanks are due to my parents, Haitao Peng and Lihua Liu, who listened to endless descriptions and various relevant thoughts of this thesis, encouraged me when I struggled or doubted myself, discussed my hypotheses and methods from their perspectives, and always supported my decisions.

TABLE OF CONTENTS

Abstract iii

Acknowledgement v

List of Figures ix

List of Tables x

Introduction 1

The Four Educational Processes 4

Before Expansion: Educational Processes in 1978-1998 7

The 1999 Higher Education Expansion 11

After Expansion: Changes in Educational Processes in 1999-2021 and Influences on Intergenerational Occupation Persistence 12

- Origin-Education Association 13
- Education-Destination Association 15
- Gaps in Destination by Origin within Education Groups 17
- Shares of Education Groups 20
- A Brief Note on Gender Disparities in Persistence 21

Analytic Approach 22

- Overall Persistence 24
- Model-Based Approach 25

Data 32

Results 35

- Overall Intergenerational Occupation Persistence 35
- Descriptive Patterns 35
- Model-Simulated Results 41
- Divergence in Descriptive and Model-Simulated Estimates 48
- Robustness Check: Taking Gender into Account 54

Discussion 55

Footnotes 62

Appendix I: Additional Tables and Figures 69

Appendix II: The Approach of Mathematical Decomposition 92

Appendix III: Additional Descriptions of Analytic Approach with Equations 109

Appendix IV: CGSS Cleaning and Processing Information 123

Appendix V: Some R Codes for Analyses 130

References 137

List of Figures

Figure 1	Percentages of Children with Bachelor's or Advanced Degrees across Parental Occupational Quartiles and Cohorts	36
Figure 2	Average Adult Occupation Rank Relative to High School Dropouts by Educational Attainment and Cohort	37
Figure 3	Rank-Rank Slopes on Occupational Attainment by Educational Attainment and Cohort	39
Figure 4	Percentages of Education Groups by Cohort	41
Figure 5	Model-Simulated Contributions of Educational Processes under Different Counterfactual Scenarios by Status of Educational Distribution	50
Figure A1	Percentages of Children with Bachelor's or Advanced Degrees across Parental Occupational Quartiles and Cohorts in Different Sub-Samples	72
Figure A2	Average Adult Occupation Rank Relative to High School Dropouts by Educational Attainment and Cohort in Different Sub-Samples	74

List of Tables

Table 1	Overall Intergenerational Occupation Persistence by Cohort and its Cross-Cohort Change in Postsocialist China	35
Table 2	Contribution of Educational Processes to Cross-Cohort in Intergenerational Occupation Persistence by Model Simulations	42
Table 3	Two-Way Interactions among Pairs of Educational Processes by Model Simulations	52
Table A1	Estimated Intergenerational Occupation Persistence by Cohort and its Cross-Cohort Change in Observation, Estimated with Age Controlled, and by Model Simulations	69
Table A2	Estimated Values: Percentages of Individuals with Bachelor's or Advanced Degrees across Parental Occupational Quartiles and Cohorts	71
Table A3	Estimated Values: Average Adult Occupation Rank Relative to High School Dropouts by Educational Attainment and Cohort	73
Table A4	Group-Specific Regression Slopes by Educational Attainment and Cohort (Age Group) and their Cross-Cohort Changes	73
Table A5	Group Shares by Educational Attainment and Cohort (Age Group) and their Cross-Cohort Changes	75
Table A6	Coefficients of Generalized Ordered Logit Regression Predicting Educational Transitions with Parental Occupation Rank	75
Table A7	Coefficients and Residual Standard Deviations of Linear Regressions Predicting Adult Occupation ISEI Scores with Parental Occupation Rank	78
Table A8	Average Parental and Adult Occupation Ranks and their Cross-Cohort Changes by Educational Attainment	82
Table A9	Model-Simulated Contributions of Educational Processes by the Status of Educational Distribution, with Statuses of Control Variables	83
Table A10	Contribution of Educational Processes to Cross-Cohort Changes in Intergenerational Occupation Persistence by Model Simulations, among Women and Men	85
Table A11	Gender Differences: Overall Intergenerational Occupation Persistence by Cohort and its Cross-Cohort Change	86

Table A12	Gender Differences: Contribution of Educational Processes to Cross-Cohort Change in Intergenerational Persistence by Model Simulations	86
Table A13	Contribution of Educational Processes to Cross-Cohort Change in Intergenerational Occupation Persistence by Model Simulations, in *Hukou*-Specific and Gender-and-*Hukou*-Specific Sub-Samples	88
Table A14	Rural-Urban (*Hukou*) Differences: Overall Intergenerational Occupation Persistence by Cohort and its Cross-Cohort Change	90
Table A15	Rural-Urban (*Hukou*) Differences: Contribution of Educational Processes to Cross-Cohort Change in Intergenerational Occupation Persistence by Model Simulations	90
Table A16	Components of Intergenerational Occupation Persistence by Educational Attainment and Cohort (Age Group) and their Cross-Cohort Changes	100
Table A17	Decomposition of Cross-Cohort Change in Intergenerational Occupation Persistence	101
Table A18	Decomposition of Cross-Cohort Change in Intergenerational Occupation Persistence, among Women and Men	105
Table A19	Decomposition of Cross-Cohort Change in Intergenerational Occupation Persistence, by *Hukou*	106
Table A20	Decomposition of Cross-Cohort Change in Intergenerational Occupation Persistence, by Gender and *Hukou*	106
Table A21	Gender Differences: Decomposition of Cross-Cohort Change in Intergenerational Occupation Persistence	107
Table A22	Rural-Urban (*Hukou*) Differences: Decomposition of Cross-Cohort Change in Intergenerational Occupation Persistence	108
Table A23	Column Names for Relevant Variables from Raw CGSS Data by Wave	129

Since the "Reform and Opening-up" starting in 1978 and continuing to the present day, China has been experiencing a multifaceted and dramatic social change (Xie 2011), which is "no less consequential for the long-term course of world history" than events commonly considered as historical watersheds in the West (Xie 2011:14-15; Xie 2012:1). In this postsocialist era, industrialization and market transition have been increasingly influential (Zhou and Xie 2019); the economic output has grown tremendously; and the country has completed the demographic transition from a high-fertility, high-mortality nation to a low-fertility, low-mortality country. In the domain of education, the scale of higher education unprecedentedly expanded in 1999, after which educational attainment in the population further improved.

Despite these achievements, with social changes between the period before educational expansion (1978-1998) and after (1999-2021), challenging new social issues have arisen. Social inequality has grown to a very high level, with the Gini coefficient in the range of 0.53~0.55 since 2005 (Xie and Zhou 2014). With such rising inequalities, intergenerational persistence of socioeconomic (dis)advantages has become an issue of great concern to both sociologists and policy makers (e.g., Chen 2013; Xie et al. 2022; Zhou and Xie 2019).

In general, the study of intergenerational persistence helps us better understand the "openness" of a society (Bloome, Dyer, and Zhou 2018) – to what extent the advantages and disadvantages in family background, such as parents' occupation status, are

reproduced and inherited. Specifically, in closed societies, the children of high-occupation-status parents become high-occupation-status adults, while the children of low-occupation-status parents become low-occupation-status adults; in open societies, by contrast, adults' occupational attainments are less constrained by their parents' occupation. Studies of intergenerational occupation persistence characterize such openness by quantifying the association between parents' occupations and their children's occupations during adulthood. Higher "persistence," i.e., more people remaining on their parents' rungs on the occupation ladder, means lower "mobility" (fewer people moving onto different rungs).[1]

In the context of China, prior research has found overall stability, with some decline, in intergenerational persistence (Chen 2013; Xie et al. 2022; Zhou and Xie 2019). However, the influence of education on the strength of intergenerational persistence and how this influence has changed over time remains to be investigated (Wu 2019). In this thesis, I first investigate the changes in intergenerational occupation persistence in postsocialist China across cohorts that came of age before or after educational expansion, exploring how societal openness has evolved. Consistent with prior research, I confirm that the intergenerational persistence remained largely stable across these cohorts.

Moving beyond prior research, I then explore the multiple roles of education in this evolution. As Bloome et al. (2018) demonstrated, the overall role of education in intergenerational persistence is made up of the contributions of four educational

processes, each of which may influence intergenerational occupation persistence. In brief, these processes are: (1) the association between parents' occupation status and children's educational attainment, (2) the correlation between children's education and occupation status in adulthood, (3) the gap in adult occupation status by parental occupation status among people with the same education, and (4) the percentage of the population in different educational groups. I assess whether the observed overall stability in intergenerational occupational persistence was due to offsetting contributions from these different educational processes, or to little change in any of the processes.

To conduct my analyses, I borrow the methodological tools of model-based simulations from Bloome et al. (2018) to evaluate the roles of each educational process. My empirical contribution, in short, is to provide a decomposition of education's multiple roles, as well as extend the estimation of persistence to younger cohorts, using the most up-to-date national survey data.

My analysis sheds light on the mechanisms of the evolution of China's social structure. In addition to affirming the dual role of education as a facilitator of both societal openness (Wu 2017) and social inequalities (Zhou 2014), this study reveals the specific elements of education that exert upward and downward pressure on the intergenerational persistence. Examining intergenerational persistence through the lens of educational processes, my thesis offers potential explanations as to why persistence and the openness of society have changed in the specific ways observed in this study.

In the following sections, I first describe the four educational processes that shape the extent of intergenerational reproduction. I then review prior studies that provide clues as to how the four processes have changed in the post-expansion period (1999-2021), compared to the pre-expansion period (1978-1998). Building on this prior research, I develop hypotheses on how these changes may have influenced intergenerational occupation persistence. I next introduce the analytic approach of model simulations, as well as the details of my data. Finally, I present, interpret, and discuss my results and discoveries.

THE FOUR EDUCATIONAL PROCESSES

Intergenerational occupation persistence refers to the extent to which the occupation status of individuals in adulthood (social destination) resembles that of their parents in the individuals' childhood (social origin) (Xie et al. 2022). To evaluate the strength of the intergenerational persistence (the degree of resemblance), we may, for example, estimate the correlation between people's social origin statuses and social destination statuses (Bloome et al. 2018:1223; Breen 2010), which captures the overall "origin-destination association."

Education plays a pivotal role in intergenerational occupation persistence (Blau and Duncan 1967; Bloome et al. 2018; Featherman and Hauser 1978; Hout and DiPrete 2006;

Torche 2015; Zhou, Moen, and Tuma 1998). As shown by Bloome et al. (2018), this role of education in overall intergenerational persistence is composed of the contribution of four processes: (1) the association between parents' occupation status and children's educational attainment (origin-education association), (2) the correlation between children's education and occupation status in adulthood (education-destination association), (3) the gaps in adult occupation status by parental occupation status among people with the same education (within-education-group gaps), and (4) the percentage of the population in different educational groups (share of education groups, or educational composition). In other words, education's contribution can be decomposed into the role these four educational processes identified above. In the following paragraphs, I introduce these four educational processes.

The origin-education association (process #1) is the correlation between one's social origin statuses and their educational attainment. Parents' occupation statuses shape their children's education; parents with better jobs are more likely to raise highly educated children (Zhou et al. 1998). With all else held constant, a higher origin-education association contributes to a higher overall origin-destination association.

Further, the education-destination association (process #2) is the correlation between an individual's educational attainment and social destination statuses. Children's education outcome predicts their adult occupation status; highly educated children typically achieve occupations that are of higher status and prestige as adults

than their less educated peers (Xie et al. 2022). All else being equal, the stronger the education-destination association, the stronger the overall origin-destination association.

Intergenerational occupation persistence is also shaped by within-education-group gaps in destination status by origin status (process #3). For example, if, among college graduates, occupational attainment tends to be higher for those whose parents had higher occupational attainment, then this will contribute to intergenerational occupational persistence, separate from the role of social origins in shaping educational outcomes.

If within-education-group gaps are larger at some education levels than others, then the shares of different education groups in the population (process #4) will also contribute to intergenerational persistence. If groups with higher within-group gaps are more prevalent, then, all else being equal, overall origin-destination persistence will be higher.

Thus, change over time in any of these four processes can adjust the overall level of intergenerational persistence in a society. Next, I discuss these four educational processes, their changes, and theoretical expectations on how the intergenerational occupation persistence has been influenced, in the social context of China. Given this thesis's investigation of people's socioeconomic attainment in the postsocialist era of China, I focus my discussion on two periods: 1978-1998, after the resumption of the *gaokao* (annual College Entrance Examination) but prior to the substantial expansion of higher education, and 1998-2021, after the expansion.

BEFORE EXPANSION
Educational Processes in 1978-1998

The resumption of the annual *gaokao* in 1977 marked the beginning of higher education development in postsocialist China.[2] Preparations for the examinations and admissions to the universities started to become standardized and nationally unified (Sun 2023). Students are admitted to universities entirely based on their rank on the exam.

Origin-Education Association

In 1978~1998, students' social origins influenced their educational opportunities and attainment. Because children whose parents held more prestigious occupations tended to score more highly in the exams (Xie 2017), it was more likely for such high-origin-status children to have more opportunities to enter universities. Besides, the resumption of *gaokao* was conducive to the emergence of shadow education in the 1980s provided by individual teachers (Feng 2021:92). Such private extracurricular tutoring services, which improved students' performances in exams (Hu, Fan, and Ding 2015), required additional family investments. Thus, students' social origin, which is associated with the plentifulness of family financial resources, influenced their educational opportunities. In short, there existed a positive correlation between one's social origin and educational attainment (origin-education association).

Education-Destination Association

Individuals' educational attainment was also positively correlated with their occupation outcomes (education-destination association) during this period, and people with university credentials tended to arrive in high-status occupations. College graduates' knowledge and skills were in high demand in urban areas to boost development (e.g., in Shanghai, see Zhou, Li, and Gao 2016). Thus, through a centralized placement system (*fenpei*), the government allocated jobs to graduates (Zhao 2016) and sent them to urban occupations in fields like "technological advancement, product upgrading, and service industry development" (Zhou et al. 2016:204). These jobs in the cities are higher status compared to the majority's jobs in agriculture (Zhou and Xie 2019).

Gaps in Destination by Origin within Education Groups

In the context of postsocialist China, existing studies have identified several factors that may lead to disparities by social origin status in people's social destination status, with both statuses measured by their occupational attainment. For instance, employees' cultural capital (Mok and Wu 2016; Scherger and Savage 2010) and social capital (Lin 2017; Mok and Wu 2016 who studied middle-class students; Yang et al. 2024 who studied rural dwellers; Zhang et al. 2010) may improve their career achievements. *Guanxi*, a unique factor in Chinese society, also influences people's outcomes (Knight and Yueh 2008; Yang and Zhang 2024). Different from social capital (Cheng 2024:327-328; Yang and

Zhang 2024:3-4), *guanxi* describes a dyadic, particular, and sentimental tie that has potential of facilitating favor exchanges between the parties connected by the tie, such as any blood or marital relationship (Bian 2006).[3] Moreover, political capital, or membership in the Communist Party of China (Nee 1989; Nee 1996), specifically, and human capital (Wu 2011; Zhuo, Ye, and Ji 2023) are also relevant.

However, in spite of these identified factors, there is little evidence specifically on the sizes of gaps in people's destination by origin *within education groups* in the pre-expansion period. Instead, what is known is that the disparities between individuals' outcomes in adulthood, as measured by their *hukou* status (Chan and Zhang 1999:822), should have been smaller among people with more education. Compared to a rural *hukou*, an urban *hukou* is associated with higher socioeconomic status (Hao, Hu, and Lo 2014; Li, Gu, and Zhang 2015) and access to high-status occupations (Wu and Treiman 2004). While there were few ways to convert *hukou* status from rural to urban under state regulations, with their successful admission to and completion of university education, college graduates with rural *hukou* were enabled to convert their *hukou* status to "urban" (Chan and Zhang 1999). As a result, they arrived in the same *hukou* status in adulthood as their urban peers, narrowing down the gaps in social status as shown by *hukou* (Wu and Treiman 2007; Wu and Zheng 2018), as well as in occupation status. In other words, in the pre-expansion period, there should have been smaller gaps in social destination by social origin among people with college degrees than their less-educated peers.

Shares of Education Groups

Meanwhile, between 1978 and 1998, higher education was the privilege of a very small number of students (Bai 2006; Li and Zhang 2010). The ratio of total enrolled undergraduate and graduate students to the total population was no larger than 0.29% (Ministry of Education 1999; National Bureau of Statistics 1999).[4] The gross enrollment ratio of higher education, which is the ratio of total enrollment in higher education regardless of age to the population of the age group that officially corresponds to higher education (UNESCO n.d.), was smaller than 6.3 percent (World Bank 2023).

Summary

In short, between 1978 and 1998 (in postsocialist China and before expansion), the origin-education and education-destination associations were positive. While there is no (to my knowledge) direct evidence on how destination gaps by social origins varied across education groups, it seems at least plausible that these within-education-group gaps were smaller for those with more education. If so, the ensuing expansion of higher education, which increased share of college graduate in the population, would reduce intergenerational occupation persistence (Duan et al. 2022). In the next section, I describe the rapid expansion of higher education in China after 1999.

THE 1999 HIGHER EDUCATION EXPANSION

In the late 1990s, Chinese government was concerned about the quantity and quality of higher education. College graduates were a small share of the employed (National Center 2001), and none of the one thousand higher education institutions could be said to be a first-class university of international recognition (Li 2004). Facing such problems, the government initiated a wave of higher education reform, growing the size of higher education scales and implementing projects to create world-class universities. Specifically, in May 1999, the authorities decided to expand the recruitment quota (the number of students to be admitted) of higher education by a big margin (Li 2004:18). Their purpose was to increase university participation rates, relieve employment pressure on secondary education graduates, and, more importantly, stimulate economic development of the education industry and other relevant fields in order to cope with the Asian Financial Crisis.

In addition, following the 1999 higher education expansion, a five-tier structure[5] gradually emerged in higher education system China (Li 2004). In general, there are 112 elite universities on the top of this hierarchy (Department of Degree Management 2008a; Department of Degree Management 2008b); the 4-year institutions, founded before and after 1999, are "ordinary" institutions in the middle; and the 3-year vocational institutions, most of whose students graduate with degrees equivalent to associate degrees, are at the very bottom.

The expansion and increasingly stratified structure of higher education, together with other social and economic changes, may have shaped intergenerational persistence in China through the four previously discussed education-related processes. In the following sections, I discuss how these four processes might have changed and influenced intergenerational persistence in the post-expansion period, compared to the earlier period.

After Expansion

Changes in Educational Processes in 1999-2021 and Influences on Intergenerational Occupation Persistence

All four educational processes identified above might have evolved in the post-expansion period, which, in turn, would alter intergenerational occupation persistence. Prior research has demonstrated that the origin-education association are stronger (Ding et al. 2021; Hu and Hibel 2014; Li 2014; Wu 2017; Wu 2013); therefore, with all else being equal, this change would increase intergenerational occupation persistence. When it comes to the education-destination association, scholars have not yet come to an agreement on how it has changed. All else being equal, if it strengthened, it would also increase persistence; on the other hand, persistence would decrease if the association weakened.

Nevertheless, prior research has not yet considered how the other two components of intergenerational reproduction – the gaps in social destination by social origin within education groups and the shares of the population belonging to different education groups – may have also contributed to changes in intergenerational persistence. To better understand the overall trend in intergenerational occupation persistence of overall stability with some decline (Chen 2013; Xie et al. 2022; Zhou and Xie 2019), in this study, I both estimate the total change in persistence between cohorts who came of age before and after educational expansion and, as my main contribution, describe how changes in the different underlying educational processes contributed to this overall change. In the sections below, I discuss relevant prior scholarship on these trends. Where there is no prior evidence, I generate theoretically grounded hypotheses for cross-cohort changes.

Origin-Education Association

Differences in students' educational attainment by their social origin increased in the post-expansion period (Ding et al. 2021; Hu and Hibel 2014; Li 2014; Wu 2017; Wu 2013), because of the increasing number of channels through which the advantages in people's family backgrounds were reflected and reproduced in their education. After the expansion, universities are allowed to charge tuition and fees (Ding et al. 2021). Such higher costs of higher education are preventing more students with disadvantaged family backgrounds from entering college (Wu 2010; Wu 2017), which further strengthened that

children from low-status families were less likely to obtain higher levels of education than their peers from high-status families.

In addition to the affordability of college education, students' advantages in family background are also facilitating their likelihood to enroll in elite universities. Students from top and locally prestigious primary and middle schools, where there are capable teaching staff, favorable learning environments, and high-achieving students (Zhou 2018), are much more likely to enter elite universities (Ye 2015). Given that access to such schools is restricted to residents in nearby neighborhoods (Chen and Park 2021; Chen, Salleh, and Binti Sohaimi 2024), parents invest in apartments in these communities to send their children to these top schools (Chen and Li 2023; Huang and Li 2014). However, most of these houses are exceptionally expensive (e.g., in Nanjing, see Xu, Song, and Liu 2018), suggesting that this pathway will only be accessible to parents with substantial economic resources in most cases.

Furthermore, while shadow education is becoming increasingly universal among students (Feng 2021; Luo, Wang, and Bai 2024; Xue and Ding 2009), wealthier families can afford more classes with higher quality (taught by more renowned and experienced instructors, with students achieving greater exam score improvements, etc.) as well as higher costs, than their poorer counterparts (Zhang and Bray 2018). Such unequal access to shadow education by social origin reproduces and reflects people's (dis)advantages in family background in their performances in exams (Hu and Hibel 2015) and ultimately

education outcome, reinforcing the association between their origin status and educational attainment as well.

Another pathway through which advantages in family background promote people's educational attainment is the alternative admission channels to (elite) universities other than *gaokao*, such as the Independent Freshman Admission Programs (IFAPs). Students selected by such programs are more likely to come from higher social origin statuses (Liu et al. 2014; Wu, Li, and Wang 2019), which exacerbates educational inequalities.

In summary, prior research has documented that the 1999 higher education expansion did not bring more equal access to educational opportunities among different social strata (Wu 2010), but contributed to the reinforcement of the origin-education association in the post-expansion period. All else being equal, this strengthened origin-education association should have increased intergenerational occupation persistence.

Education-Destination Association

When it comes to the association between people's educational attainment and social destination status, prior research has not yet reached an agreement on how it has changed in the post-expansion period. While it is clear that the economic returns to high school credentials (increases in income due to completing high school education, compared to high school dropouts) has remained stable (Hu 2013), there is evidence

suggesting both a weakened (e.g., Yang 2024) and a strengthened (e.g., Hu and Hibel 2015) education-destination in later compared to earlier cohorts.

Taking the sectors of employment (government vs. market) into account, Yang (2024) discovered that the higher education expansion weakened the connections between college credentials and high-status occupations, because there was an increasing number of college graduates entering non-government sectors of lower-status jobs. Her discovery indicates a decreased education-destination association. On the other hand, Hu and Hibel (2015) suggest higher economic returns to higher education, that college graduates are benefiting more from their college degrees and enjoying a higher increase in their income, in the post-expansion period. Their analyses imply that the association between people's education outcome and social destination has otherwise increased.

There are also disagreements when more details are involved, taking the tiers of universities and levels of higher education into account. As the market sectors emerge during the market transition (Li 2003; Wu and Xie 2003; Yang 2005), employers who provide high-status occupations and internationally competitive salaries usually prefer graduates from (elite) universities (Chan and Zhang 2021). In other words, individuals with degree(s) from (elite) universities are more likely to arrive in higher social destination statuses. As a result, in the post-expansion period, the returns to education of all college graduates declined, except those with elite university degrees for whom the return remained high (Kang, Peng, and Zhu 2021). The returns to education may also not

change monotonically over time. From the pre- to the post-expansion period, while the returns to advanced degrees (master's or above) kept increasing, the return to undergraduate education increased first, then decreased (Li, Wu, and Meng 2022). Together, these mixed findings indicate that post-expansion changes in returns to education, and therefore the education-destination association, are still unclear. Therefore, rather than developing specific hypotheses on the changes in these associations, I treat their change as an empirical question and one contribution of my analyses.

Gaps in Destination by Origin within Education Groups

Strengthened origin-education and education-destination associations alter persistence via differences *between* education groups (in terms of parental and adult occupation status). What about changes *within* education groups? I hypothesize that, in the post-expansion period, the adult outcome gaps by origin status among college graduates have enlarged, while the gaps may have remained stable among high school dropouts and graduates.

In postsocialist China, workers with and without college education occupy two different segments of the labor market, respectively, with different driving forces of income growth (Wu 2011). As a result, the contributors of occupational attainment among people with more education tend to be different from those of their less-educated

counterparts. I therefore develop separate predictions for these two education groups about how within-group disparities in outcomes by social origins may have changed over time.

In the post-expansion period, Wu (2011) discovered that, as they age, human capital accumulation (education and seniority, including total experience and firm-specific job tenure) is the most important factor in boosting the growth of college graduates' occupation status, while job mobility (whether they have changed jobs and how often) has no effect. On the other hand, people with less education move upwards on the income ladder not because of their human capital accumulation but job mobility.

For people with college degrees, I hypothesize that gaps in occupation outcomes by origin status have grown. Even though the education outcomes of individuals with college degrees are quantitively similar, their educational attainment is qualitatively different by social origin – people graduate from different tiers of higher education institutions (Li 2004; Yang and Wang 2020), and these tiers have become increasingly stratified by students' social origins (Ding et al. 2021; Du and Li 2024; Liu et al. 2014; Wu 2019). Given that education outcomes are one way by which scholars measure human capital (Wu 2011:74), I hypothesize that such growing qualitative disparities in people's educational attainment by their social origin (Chan and Zhang 2021; Ding et al. 2021) produce larger gaps in people's destination status by origin status. Prior research has also

found larger gaps in social destinations by social origin among college graduates in the post-expansion period (Duan et al. 2022).

On the other hand, there is little evidence that the within-education-group gaps in social destination by social origin among high school dropouts and graduates have changed. As noted above, job mobility, not human capital, drives the occupational attainment of less-educated people but job mobility (Wu 2011). However, though there have been higher rates of job mobility in the post-expansion period (Li 2013), no prior research has discovered more rapid increases in job mobility for those from more advantaged social origins. In addition, I am aware of no findings on differential changes by social origins in cultural capital, social capital, *guanxi*, and CPC membership (political capital) among those with less education either. Therefore, given that there is no evidence in support of changes in the within-education-group gaps among high school dropouts and graduates to the best of my knowledge, I hypothesize that their gaps have remained stable.

In conclusion, I hypothesize that, after the 1999 higher education expansion, the within-education-group gaps in the occupation status by origin status of college graduates have enlarged, while those of their less-educated counterparts generally remained stable. Overall, then, I hypothesize that changes in within-education disparities in outcomes by social origins strengthened intergenerational persistence. One contribution of this thesis is to estimate how changes in within-education-group gaps in

occupational attainment by social origins contributed to trends in intergenerational occupational persistence.

Shares of Education Groups

A fourth potential contributor to changing intergenerational occupation persistence is the changing educational composition of the population. For postsocialist China, specifically, I hypothesize that the increased share of college graduates after the educational expansion have reduced persistence. More people completing college has the potential to disrupt intergenerational persistence because the origin-destination association is weaker among college graduates than among non-graduates (Wu et al. 2019; Yu et al. 2023). Thus, as more people shifted into a portion of the educational distribution where the origin-destination association is weaker, overall intergenerational persistence would decline (Breen 2010:368).

The scale of higher education in postsocialist China substantially enlarged after the 1999 higher education expansion. From 1998 to 2022, both the ratio of total enrolled college students to the population[6] and the gross enrollment ratio of higher education (the ratio of total enrollment in higher education regardless of age to the population of the age group that officially corresponds to higher education, UNESCO n.d.) multiplied tenfold (Ministry of Education 1999; Ministry of Education 2023; National Bureau of Statistics 1999; National Bureau of Statistics 2023). In addition, the number of new

graduates expanded more than sixfold, increasing from 0.8 million in 1998 to over 5.5 million in 2022.[7] The expansion indeed allows more youths to attain college education(s) than before (Yeung 2013). As a result, the proportion of laborers with college degrees in the younger cohorts has also increased.

In summary, considering the theory of Breen (2010) and Hout (1988), the increased share of college graduates after the expansion of higher education in 1999 should have reduced the persistence, because of the weaker origin-destination association among college graduates. In my analyses, I explore how these variations, as well as changes in the origin-education association, education-destination association, and gaps within education groups, contribute to trends in intergenerational occupation persistence in postsocialist China.

A Brief Note on Gender Disparities in Persistence

Given that women's employment is nearly universal throughout the history of the People's Republic of China (Xie 2014), scholars studying intergenerational persistence in China have also explored disparities in persistence by gender (e.g., Huo and Golley 2022; Xie et al. 2022). Xie et al. (2022) discovered that women tended to experience (much) higher intergenerational persistence than men in the same cohorts.

In addition, women and men have also experienced different changes in some aspects of their education. On one hand, while the association between people's social

origin status and their educational attainment has strengthened for both genders, the gain has been larger for women than for men (Huo and Golley 2022). On the other hand, the substantial expansion in higher education has significantly decreased inequalities in access to higher education between women and men (Song and Tan 2024; Wu and Du 2018). The gross enrollment ratio of higher education (total enrollment in higher education regardless of age out of the population of the age group that officially corresponds to higher education, UNESCO n.d.) of women even exceeded that of men in the past decade (World Bank Gender Data 2023).

Therefore, changes in the four educational processes – the origin-education association, education-destination association, gaps within education groups, and shares of education groups – may have influenced persistence trends differently for women and for men. While these potential gender disparities are not the primary focus of my analyses, I report and briefly interpret these differences after discussing my main empirical results. In the next sections, I introduce my analytic approach and data in detail.

Analytic Approach

In this study, there are three key analytical variables – individual's social origin status, educational attainment, and social destination status. The social origin and destination statuses are measured with the parental and adult occupational percentile

ranks, following practice in prior research (e.g., Xie et al. 2022). With these ranks and information on educational attainment for cohorts who came of age prior to the educational expansion (1964-1979) and after (1980-1995), I first report the descriptive patterns of the association between parental occupation and offspring education (origin-education association, process #1), and then between offspring's education and occupation in adulthood (education-destination association, process #2). I also describe the gaps in social destination by social origin within education groups (process #3), as well as the shares of each group (process #4), across cohorts. These descriptive patterns are helpful for developing a basic sense of how the four educational processes have changed in the post-expansion period, and therefore their potential contributions to the intergenerational occupation persistence.

Next, I estimate the overall intergenerational occupation persistence, using a linear regression that predicts adult occupational rank (social destination status) with parental occupational rank (social origin status) for each cohort and compare results across cohorts. This tests how overall intergenerational persistence has evolved (strengthened, stayed stable, or weakened) in the more recent cohorts.

I then proceed with the core contribution of my analysis – to explore the multiple roles of education in intergenerational occupation persistence. Specifically, I borrow the model-based approach introduced by Bloome et al. (2018), to show how changes in (1) origin-education association, (2) education-destination association, (3) within-education-

group disparities in outcomes by social origins, and (4) the educational distribution are estimated to have jointly contributed to the observed changes in overall persistence. I discuss the estimation of overall persistence and the model-based approach in more detail below. (An alternative approach to explore the multifaceted role of education is the "mathematical decomposition"; see Appendix II for details.)

Overall Persistence

After reporting the descriptive patterns, I estimate overall intergenerational occupation persistence and its change across cohorts. For each cohort of interest, intergenerational occupation persistence can be measured as the slope coefficient from a linear regression of adult occupational percentile rank on parental occupational percentile rank[8],

$$O_i^a = \alpha + \beta \cdot O_i^p + \epsilon_i \quad (1)$$

with O_i^a being the adult occupational rank of observation i, their parental occupational rank recorded as O_i^p, and ϵ_i being the error. Thus, β (the slope coefficient) captures persistence, and the complement of this slope coefficient, $1 - \beta$, captures mobility. To estimate the difference in persistence across cohorts, I calculate $\beta^{young} - \beta^{old}$, where β^{old} and β^{young} are the slope coefficients for people in the older cohort (1964-1979) and those in the younger cohort (1980~1995), respectively.

Using percentile ranks of parents' and children's occupation statuses within their respective group's distributions, β can be further interpreted as an intergenerational rank-rank correlation that captures purely relative persistence. Sociologists have long prioritized understanding patterns in purely relative persistence, explicitly conditional upon marginal class distributions (Breen 2004; Featherman and Hauser 1978; Wang 2024:129-130). Excluding the influences of structural changes in society, including in the labor market, relative persistence reveals the extent to which the offsprings' positions in their occupation distribution resembles the parents' position in parents' occupation distribution. Thus, when studying postsocialist China, a society that has been experiencing structural changes in its labor market (Cai and Wang 2010; Zhou 2014), it is valuable to focus on relative persistence, to describe the persistence experienced by individuals with various positions in their respective distributions.

Model-Based Approach

Next, to understand the role of education in the persistence trend, I borrow the methodological tool of model-based simulations from Bloome et al. (2018:1224-1226). There are three steps in total. First, I estimate four sets of parameters, each specifically capturing one educational process, separately for the older cohort (1964-1979, before educational expansion) and younger cohort (1980-1995, after expansion). Then, with these parameters, I simulate the intergenerational occupation persistence of each cohort,

as well as the cross-cohort change. Such a framework of simulation (with the four individual sets of parameters) helps to identify the specific contribution of each educational processes to the cross-cohort change in intergenerational persistence: I can manipulate the status (changed as observed or not) of each of the four processes separately to develop counterfactual scenarios, in which any one or more of the processes did *not* change as observed but stayed unchanged, to isolate the role of each process. Therefore, the last step is to conduct these counterfactual analyses to isolate the individual roles of each educational process. I introduce these three steps in detail below. Sampling uncertainty is assessed with 95% intervals from repeated simulations.[9]

The Models

For the analyses isolating the contribution of each of the four educational processes, I estimate two different models. The first predicts educational attainment using a generalized ordered logit model (Peterson and Harrell 1990; Williams 2006; Williams 2016),

$$\text{logit}[P(E_i > g)] = \alpha_g + \lambda_g \cdot \left(O_i^p - \mu^p\right) \quad (2)$$

where, for observation i, E is educational attainment, O^p is parental occupational rank, μ^p is the population mean of parental rank, and α_g and λ_g are the intercept and slope coefficients associated with attaining education level g. I recenter parental occupational rank at its population mean. Therefore, changes in α_g across cohorts capture changes in

educational composition, and cross-cohort changes in λ_g capture changes in the strength of the origin-education association.

The second model is to predict adult occupation attainment using the education levels estimated from the first model. Specifically, a normal linear model is used to predict the adult occupation ISEI score, I_i^{sa}, within each education group using parental occupation rank. As shown in Equation 3, the predicted adult ISEI, I_i^{sa}, is assumed to follow a normal distribution where the mean is $\gamma_g + \beta_g \cdot \left(O_i^p - \mu^p\right)$ and standard deviation is v_g. Again, I recenter parental occupational rank at its population mean.

$$I_i^{sa} \sim N\left[\gamma_g + \beta_g \cdot \left(O_i^p - \mu^p\right), v_g^2\right] \tag{3}$$

Therefore, cohort changes in γ_g capture changes in the education-destination association for people with median parental occupation status; changes in β_g and v_g capture the changed gaps in destination status by origin status within education groups.

In addition to facilitating counterfactual analyses (more details below), by capturing the educational processes with four individual sets of parameters in the two models, another strength of the models is to allow changes in one educational process to alter key quantities relevant to other educational processes. For example, changes in the shares of education groups (captured by α_g) and in the origin-education association (captured by λ_g) are allowed to adjust the means and variances of parental occupational ranks within education groups.

Furthermore, while they follow the footsteps of the status attainment model introduced by Blau and Duncan (1967), the two models also relax the traditional assumptions of linearity and additivity. The assumption of linearity is relaxed with educational attainment modeled as an ordered, categorical outcome (in Equation 2) rather than a continuous variable (e.g., years of schooling). The assumption of additivity is relaxed because the full interaction effect of parental occupation status and children's education on adult occupation status is allowed (via separating β_g and υ_g parameters in each education group, in Equation 3).

At this step, operationally, I estimate all the parameters in Equations 2 and 3 ($\hat{\lambda}_g$ for origin-education association, $\hat{\gamma}_g$ for education-destination association, $\hat{\beta}_g$ and $\hat{\upsilon}_g$ for within-education-group gaps, and $\hat{\alpha}_g$ for educational composition) of the older cohort and of the younger cohort, with their respective sample data. The next step is to simulate the persistence in each cohort.

Persistence Simulation

When simulating persistence, the parameters of the generalized ordered logit model (Equation 2) predict the individuals' probability distributions of the levels of education. Suppose there are n levels of education in total, i.e., the education group g takes on the values $1, 2, \ldots, n$, the model can be written as (Bloome et al. 2018; Williams 2006; Williams 2016)

$$P(E_i > g) = \frac{\exp\left[\alpha_g + \lambda_g \cdot \left(O_i^p - \mu^p\right)\right]}{1 + \left\{\exp\left[\alpha_g + \lambda_g \cdot \left(O_i^p - \mu^p\right)\right]\right\}} \qquad g = 1, 2, \dots, n-1 \tag{4}$$

where E_i is the educational attainment of observation i, O_i^p is their parental occupational rank, μ^p is the population mean of parental rank, and n is the total number of levels of education. Then, according to Equation 4, the probabilities of respondent i attaining each level of education are equal to

$$\begin{aligned} P(E_i = 1) &= 1 - P(E_i > 1) \\ P(E_i = g) &= P(E_i > g - 1) - P(E_i > g) \qquad g = 2, \dots, n-1 \\ P(E_i = n) &= P(E_i > n - 1) \end{aligned} \tag{5}$$

Based on these probabilities, I predict the level of education that each respondent would attain, and therefore the distribution of educational groups.

Next, conditional upon the level of education simulated with Equations 4 and 5 above, the parameters of the normal linear model (Equation 3) predict the adult occupation ISEI score. When the observation i, whose parents' occupational rank is O_i^p, attains educational level g, their expected adult ISEI, I_i^{sa}, is calculated according to

$$I_i^{sa} = \gamma_g + \beta_g \cdot \left(O_i^p - \mu^p\right) + z_i^* \cdot v_g \tag{6}$$

where the multiplier of randomness, z_i^*, is the inverse of the standard normal cumulative distribution function, evaluated at a randomly generated value under a standard uniform distribution for observation i. This multiplier of randomness reflects the uniqueness and randomness of individuals' growths – not all individuals with the same social origin and educational attainment attain occupation outcomes of the same rank. After that, I

transform the simulated adult ISEI into percentile ranks. (More details in Appendix III – The Model-Based Approach – Persistence Simulation.)

Lastly, the simulated intergenerational occupation persistence is the slope coefficient of the linear regression simulating the expected adult occupational rank, O_i^{sa}, with the observed parental occupational rank, O_i^p, as shown in Equation 7 below.

$$O_i^{sa} = \alpha + \beta^s \cdot O_i^p + \epsilon_i \tag{7}$$

It is exactly the same as Equation 1, except that the regression predicts the simulated adult occupational ranks (O_i^{sa}), instead of the observed adult ranks (O_i^a). Therefore, the regression slope coefficient in Equation 14, β^s, captures the simulated intergenerational occupation persistence, whose superscript s denotes "simulation."

Counterfactual Analyses

To explore the specific roles of each educational process in the cross-cohort changes of the intergenerational occupation persistence, I rely on counterfactual analyses, targeting explicitly one process at a time. I first simulate what the persistence would have been if the targeted educational process had changed as observed. Next, with the other three educational processes in the same statuses (changed or unchanged) as in the first scenario, I simulate what the persistence would have been if the process of interest had *not* changed but remained stable. In this pair of counterfactual scenarios, two persistence values are simulated with only the status of the targeted educational process being

different, with those of the other three processes under control. Finally, I subtract the second persistence value (targeted educational process *un*changed) from the first value (targeted process changed) to get the contrast. This contrast determines the role of the educational process of interest: if the difference is positive, then the process is predicted to have increased intergenerational occupation persistence in the post-expansion period, and vice versa.

Specifically, to simulate the counterfactual scenarios where the educational processes changed as observed, I use the younger cohorts' parameters (λ_g^{young} for origin-education association, γ_g^{young} for education-destination association, β_g^{young} and υ_g^{young} for within-education-group gaps[10], and α_g^{young} for educational composition). On the other hand, the parameters of the older cohort (denoted by superscript *old*) simulate what would have happened if the educational processes had *not* changed. In other words, to determine the roles of, for example, origin-education association, I rely on the contrast of persistence simulated in such a pair of counterfactual scenarios: one scenario involves λ_g^{young} (changed as observed) and the other λ_g^{old} (did *not* change), with the statuses of other three educational processes being the same in both scenarios.

In addition, for each of the four educational processes, there are eight such pairs[11] of counterfactual scenarios. Two of these eight pairs are special, one with the other three processes all changed (using younger cohorts' parameters) and the other with all other three remained unchanged (using older cohorts' parameters), respectively. Suppose our

targeted process of interest is, for example, the origin-education association (captured by λ_g). In the former pair of comparison, λ_g is the only unchanged parameter and all others are changed. Bloome et al. (2018) used this measure to capture the role of origin-education association (λ_g), estimating how different intergenerational persistence would have been for the younger cohort, had this parameter not changed (λ_g), compared to its simulated outcome with all parameters changed as observed. On the other hand, the latter pair of comparison captures the role of origin-education association (λ_g) by indicating how persistence for the younger cohort would have differed from its counterfactual outcome with all parameters unchanged[12], had only λ_g changed and all others remained unchanged. Moreover, the average of all eight contrasts from these pairs of counterfactual scenarios also evaluates the roles of educational processes (see Appendix III – The Model-Based Approach – Counterfactual Analyses for details). In my analyses, I report estimates from all these three ways of evaluation.

DATA

To study the multiple roles of education in intergenerational occupation persistence in postsocialist China, I use data from nine waves of the Chinese General Social Survey (CGSS) conducted between 2008 and 2021. CGSS is a nationally representative, repeated, and cross-sectional survey that has been conducted by the National Survey Research

Center at Renmin University of China since 2003 (Bian and Li 2012; Xie et al. 2022; Zhou and Xie 2019). In this study, I pool the nine samples to form a single datafile by extracting information on year of birth, age, current job, the most up-to-date completed highest level of education at the time of the survey[13], the parents' occupation when the respondent was 14, survey weights, as well as gender and *hukou* (household registration) status.

Earlier research has shown that, in postsocialist China, young workers are most likely to change their jobs and switch to new employers (Whyte and Parish 1984; Zhou, Tuma, and Moen 1997). Thus, early-adulthood measures of occupational attainment are weaker proxies of adult lifetime attainment, tending to understate the levels of persistence and thus providing biased estimates. To minimize such life cycle effects that might confound observed trends across generations, I therefore restrict the respondents in the sample to those who were 25 or older at the time of the survey.

To estimate the overall intergenerational occupation persistence before and after the 1999 higher education expansion and contributions of changes in the four educational processes, I compare two cohorts (age groups): people born between 1964 and 1979 (older cohort)[14] and those born from 1980 to 1995 (younger cohort)[15]. Both cohorts consist of a span of 16 years in total, and all people's social origins were completely situated in postsocialist China.[16] However, the 1999 higher education expansion only influenced the university admission, experiences, and achievements of the people in the younger cohort,

who were 18 or younger in 1999. Thus, comparing the intergenerational occupation persistence of these two cohorts enables me to compare the roles of education in persistence before and after the educational expansion.

I determine the percentile ranks of the parental and adult occupation statuses according to the International Socio-Economic Index (ISEI, see Ganzeboom, De Graaf, and Treiman 1992) scores of their jobs that are recorded in the CGSS with the 1988 International Standard Classification of Occupations (ISCO-88, see Ganzeboom and Treiman 1996) codes.[17, 18, 19] For simplicity, I take the higher value between the father's and the mother's ISEI scores to indicate the parental occupation status. The actual percentile rank of each ISEI score is then taken as the mid-point value of its percentile range within each cohort.[20] My analytic sample sizes are 20,626 and 8,468 for the older and younger cohorts, respectively.

To assess education's multiple roles in intergenerational occupation persistence, I classify people into four education groups according to their highest level of education: less than high school ("<HS"), high school ("HS"), junior college degree ("AA")[21], and bachelor's degree or above ("BA+")[22, 23].

(For details of how CGSS raw data are cleaned and processed, see Appendix IV; for how the percentile ranks are assigned based on mid-point value of percentile ranges, see Appendix V – Percentile Rank.)

RESULTS

Overall Intergenerational Occupation Persistence

Table 1 records the estimated rank-rank slopes on occupational attainment for the two cohorts, which indicate the overall intergenerational occupation persistence. In postsocialist China, intergenerational persistence remained stable between the cohorts – it was trivially and statistically insignificantly lower in the younger cohort (.310) than in the older cohort (.335) – consistent with prior empirical results (Chen 2013; Xie et al. 2022; Zhou and Xie 2019). Such sizes of intergenerational rank-rank correlations indicate a moderate level of occupation persistence and, on the flip side of the coin, a moderate amount of mobility.

Table 1. Overall Intergenerational Occupation Persistence by Cohort and its Cross-Cohort Change in Postsocialist China

	Older Cohort (1964-1979)	Younger Cohort (1980-1995)	Δ $\beta^{young} - \beta^{old}$
Rank-Rank Slope	.335	.310	−.025
	(.320, .350)	(.286, .334)	(−.053, +.004)

Source: CGSS 2008~2021 Data.

Note: Bootstrapped 95% intervals in parentheses.

Descriptive Patterns

Origin-Education Association

Before showing the results of model-based simulations, I describe trends in the four educational processes over time. Figure 1 shows the percentages of offspring receiving a

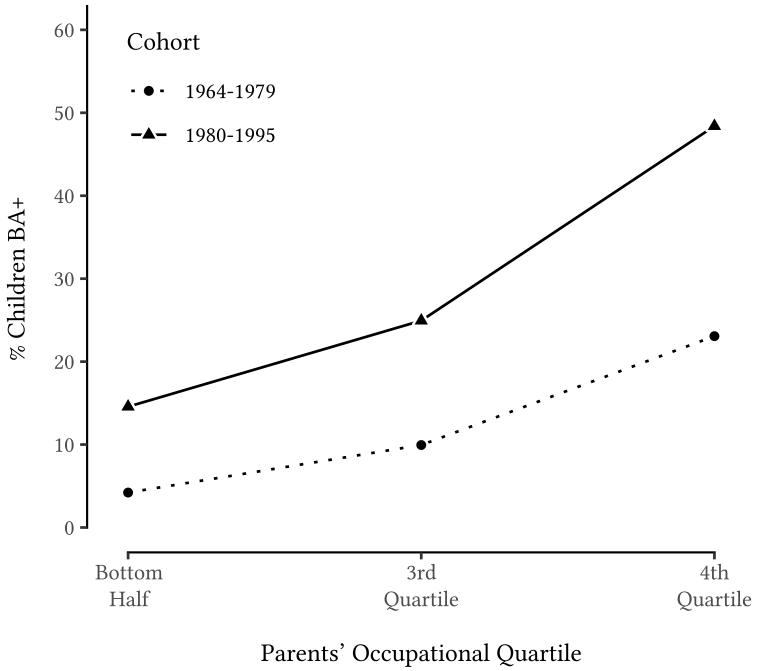


Figure 1. Percentages of Children with Bachelor's or Advanced Degrees across Parental Occupational Quartiles and Cohorts

Source: CGSS 2008~2021 Data.

Note: The 1st and 2nd Quartiles of parents' occupation percentile ranks are combined as the "Bottom Half" category. Estimated values are in Appendix I – Table A2.

bachelor's degree or above across cohorts, stratified by social origin statuses.[24] As expected, social origins are positively associated with educational attainment in both cohorts, and the association strengthened after the expansion. In the older cohort, people from the bottom half of the parental occupation distribution were about 19 percentage points less likely to complete college than people from the highest quartile. In the younger cohort, this gap had grown to over 33 percentage points. This strengthened origin-education association (process #1) in the post-expansion period likely contributed to increase the overall intergenerational occupation persistence.

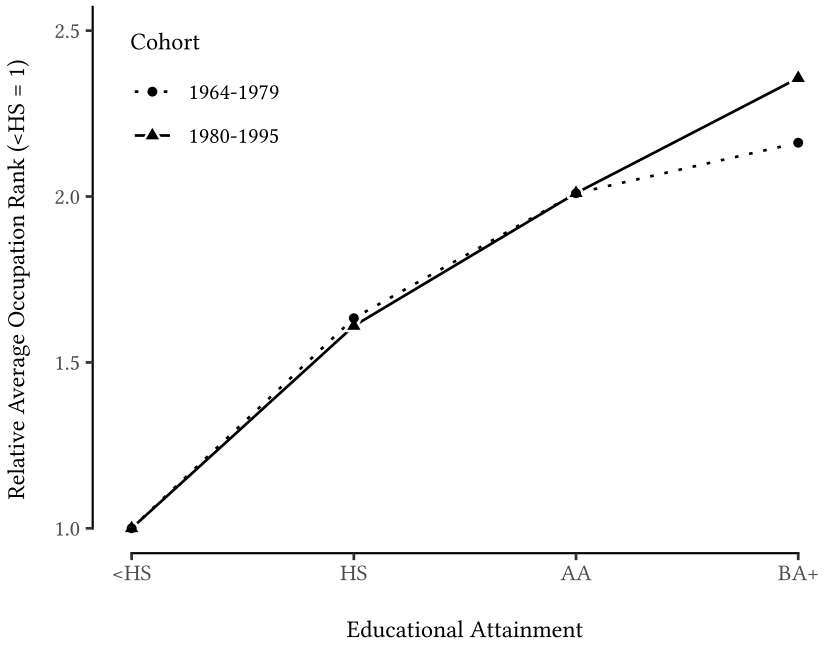


Figure 2. Average Adult Occupation Rank Relative to High School Dropouts by Educational Attainment and Cohort

Source: CGSS 2008~2021 Data.

Note: High school dropouts (<HS) are individuals without a high school diploma, and their relative rank (to themselves) is set to 1. Estimated values are in Appendix I – Table A3.

Education-Destination Association

The correlation between people's education and destination status was also positive in both cohorts and only somewhat stronger in the younger cohort. Figure 2 shows the average occupational ranks in people's adulthood across the four education groups, relative to that of high school dropouts, in the two cohorts. Therefore, the numerical value of one group describes how many times higher its average rank is than that of their counterparts without high school diplomas. Stratification across education groups was similar across cohorts for those without a college degree. Specifically, in both cohorts, a

high school diploma increased people's average occupation rank by roughly 60 percent compared to high school dropouts, while the improvements from an associate's degree were slightly over 100 percent. On the other hand, only the advantage of a bachelor's degree or above increased across cohorts. In the older cohort, the average occupational rank of those with a bachelor's degree was about 120 percent higher than that of high school dropouts. In the younger cohort, it was nearly 140 percent larger. Therefore, this "somewhat" reinforced education-destination association (process #2) likely increased intergenerational occupation persistence, but perhaps not to a large extent.

Gaps in Destination by Origin within Education Groups

Next, Figure 3 shows the group-specific rank-rank slopes of each education group across cohorts. I applied the same method as that for the overall persistence, except that the estimations were on education-specific samples.[25] Therefore, the estimates reflect how strongly occupational rank persists across generations in each education group. If the solid line (for the younger cohort) is steeper than the dotted line (for the older cohort), there is a growing gap in an education group, which represents a stronger intergenerational occupation persistence and an increasingly decisive role for parental occupation status in people's occupation in adulthood, among individuals with the given level of education. All else being equal, expanding gaps within education groups tend to raise the entire population's persistence.

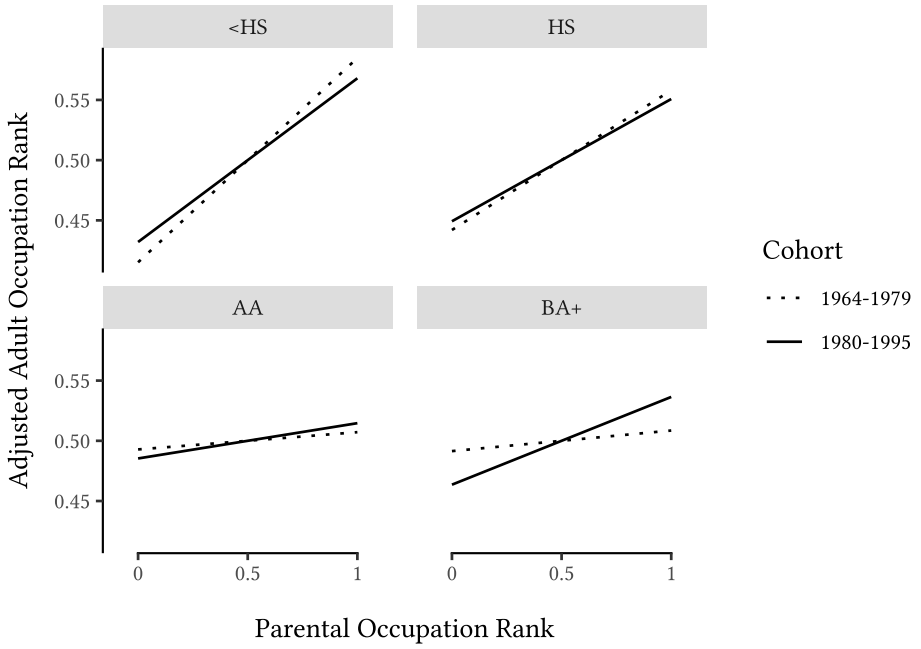


Figure 3. Rank-Rank Slopes on Occupational Attainment by Educational Attainment and Cohort

Source: CGSS 2008~2021 Data.

Note: Steeper lines reflect higher rank-rank slopes. The estimated adult ranks are manually adjusted such that the adult rank associated with 50th percentile parental rank is also 50th percentile. Therefore, this figure does not reflect the actual y-intercept (expected adult occupation rank given that the parental occupation rank is 0th percentile). Estimated values are in Appendix I – Table A4.

On one hand, different from my hypothesis, within-education-group gaps among less-educated individuals actually decreased in the post expansion period (less steeper solid lines). The within-education-group slope declined among high school dropouts (from .170 to .136) and high school graduates (from .116 to .102). Therefore, in the post-expansion period, among those with less education, parental occupation played a smaller role in determining occupational outcomes, which tended to weaken the overall persistence.

On the other hand, I hypothesized that the within-education-group gaps for people with higher education credentials increased. In support of my hypotheses, the gaps grew for both people with an associate's degree (from .014 to .029) and those with a bachelor's degrees or above (from .017 to .073). In other words, the level of persistence within groups of higher education strengthened, tending to strengthen persistence in the entire population. In spite of these increases, within-group persistence is still lower among those with higher education than among their less-educated peers.

Therefore, when it comes to the changes in gaps within education groups (process #3), there are disparities at different parts of the educational distribution. Given these countervailing trends, it is difficult to tell how they will influence intergenerational occupation persistence and further efforts are necessary.

Shares of Education Groups

Lastly, the shares of education groups (process #4) are recorded in Figure 4. In the post-expansion period, the percentages of high school dropouts decreased to only 40 percent from nearly 70 percent, while the share of high school graduates increased by approximately 5 percent, and that of associate's degree holders nearly doubled (from 8 percent to 16 percent). The substantial educational expansion mainly happened at the level of bachelor's degree or above, whose percentage roughly tripled from 9 percent to 25 percent. In other words, the 1999 higher education expansion has shifted more people

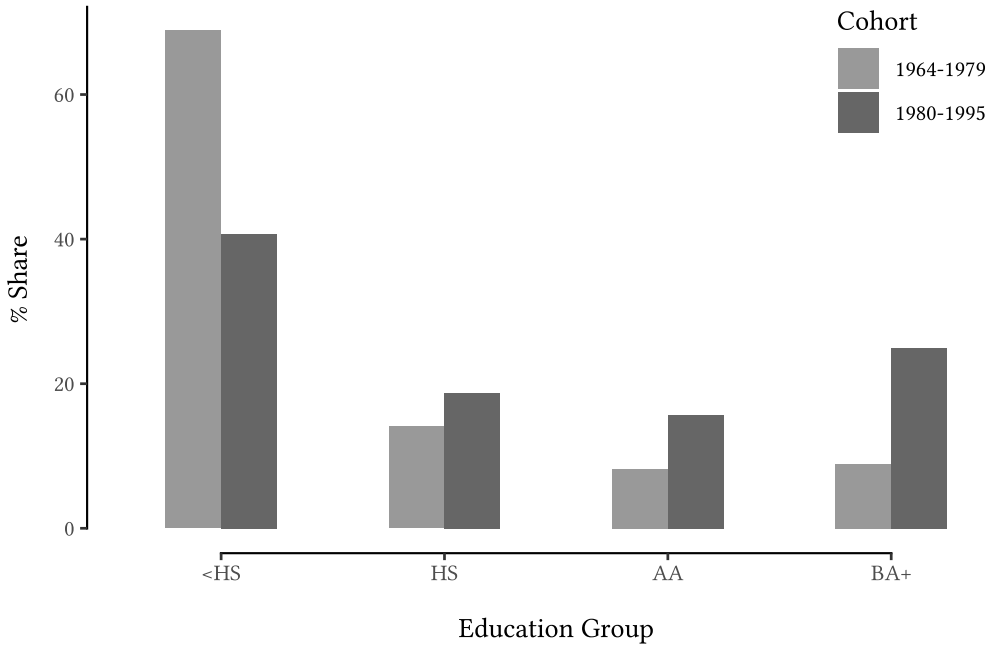


Figure 4. Percentages of Education Groups by Cohort

Source: CGSS 2008~2021 Data.

Note: Estimated values are in Appendix I – Table A5.

to higher levels of education, where there are relatively lower levels of persistence (see the section above). Therefore, all else being equal, such an educational upgrading contributed to a decline in the overall intergenerational occupation persistence in the post-expansion period.

Model-Simulated Results

I formally estimate the contribution of these four educational processes to the cross-cohort change in the intergenerational occupation persistence with the models. The model-simulated results are close to the observed cross-cohort changes in persistence

(see Appendix I – Table A1), so it is confirmed that the steps of simulation (Equations 2~7) are able to accurately reproduce the estimates of intergenerational occupation persistence.

Table 2. Contribution of Educational Processes to Cross-Cohort Change in Intergenerational Occupation Persistence by Model Simulations

	As Only Unchanged	As Only Changed	Full Average of 8 Pairs
Origin-Education Association	–.024	–.020	–.022
	(–.057, +.010)	(–.054, +.013)	(–.034, –.010)
Education-Destination Association	–.032	–.031	–.032
	(–.065, +.001)	(–.065, +.002)	(–.044, –.020)
Within-Education-Group Gaps	+.019	–.005	+.007
	(–.014, +.053)	(–.038, +.029)	(–.005, +.019)
Shares of Education Group	+.028	+.019	+.023
	(–.006, +.062)	(–.014, +.051)	(+.011, +.035)

Source: CGSS 2008~2021 Data.

Note: 95% intervals from repeated simulations in parentheses. Column 1 treats the target factor as the only unchanged (using old parameters), while other three changed as observed. Column 2 treats the target factor as the only changed (using young parameters). Column 3 takes the average of all 8 possible pairs of comparison. See Appendix III – The Model-Based Approach – Counterfactual Analyses for details.

The results of model-based counterfactual simulations are shown in Table 2, with three ways to evaluate the contribution of educational processes to overall persistence. Essentially, I measure what the persistence would be if the educational process of interest changed, compared to if it did *not* change, while holding the other three processes controlled. The first set of comparisons allows the other three processes to all change; the second allows none of the other three processes to change; the third considers all eight possible statuses of the other three processes. Because the third way involves

multiple pairs of comparison, it tends to generate narrower confidence intervals. The greater precision of the third way does not mean it is theoretically and conceptually more convincing, merely that it uses information from more pairs of comparisons. Therefore, I look at the results in all three columns, and they show offsetting influences across the four educational processes.

Origin-Education Association

The first row of Table 2 investigates the contribution of the association between parents' occupation status and children's education (origin-education association, process #1). The point estimates from all three ways of evaluation suggest that the origin-education association decreased overall intergenerational persistence, and the association is statistically significant in the third way of evaluation. This result is contrary to both my theoretical expectation and the descriptive pattern in Figure 1.

What explains this divergence? Recall that the outcome of target is children's educational attainment. From the perspective of the models (Equation 2), the distribution of individuals from different social origin statuses into educational categories is shaped by two educational processes, the association between parents' occupation status and children's educational attainment (captured by λ_g) and the educational distribution (captured by α_g). Changes in either of these two processes (λ_g and/or α_g) influence individuals' educational attainment. In China's post-expansion period, both λ_g and α_g

evolved. Therefore, for the role of origin-education association (process #1) to the cross-cohort change in persistence demonstrated by the descriptive pattern (Figure 1), there are two sources – a). the "true changes" in origin-education association (λ_g) and b). some additional influences relevant to the educational upgrading (changed α_g). In other words, the descriptive pattern did not show the pure contribution of our process #1, origin-education association.

Sample statistics identified one of the "additional influences relevant to α_g" as the movement of the overall population mean. Before educational expansion, almost 70 percent of the cohort lacked a high school degree (Figure 4 and Appendix I – Table A5). The overall population mean of parental occupation rank (.500) was somewhere between the group average of high school dropouts (<HS, .440) and high school graduates (HS, .584) (Appendix I – Table A8). When almost 60 percent of the cohort had at least a high school degree after the expansion, however, the group average of high school graduates dropped below the overall population mean. Therefore, in the post-expansion period, the overall population mean relocated to somewhere between high school graduates and people with an associate's degree (AA). In other words, the overall population mean "moved upwards."

In short, there were two sources in the role of origin-education association (process #1) demonstrated by the descriptive pattern (Figure 1) – the true contribution (due to changes in λ_g) and some additional influences following the educational upgrading (the upward movement of the population mean, due to changed α_g). While the descriptive

pattern could not detangle the two, the models can identify and isolate the explicit, pure role of the origin-education association. This strength of the model simulations resulted in the divergent descriptive and model-simulated results.

Education-Destination Association

Similarly, changes in the association between children's education and occupation status in adulthood (education-destination association, process #2), shown in the second row of Table 2, also contributed to decreasing intergenerational persistence (statistically significant when the full set of comparisons is used). This result is contrary to my descriptive patterns in Figure 2. The divergence, again, resulted from the fact that Figure 2 could not capture the pure contribution of education-destination association.

From the perspective of models, educational upgrading still brings additional influences on intergenerational persistence, in addition to the true contribution of education-destination association, though implicitly. Recall that Equation 2 predicts the individuals' educational attainment, and the prediction is influenced by α_g that captures educational upgrading. Further recall that Equation 3 that predicts children's occupation in adulthood is conditional upon Equation 2 – I determine which sets of parameters to use in Equation 3, based on the predicted level of education in Equation 2. For example, if one is predicted to attain a high school degree by Equation 2, I use the parameters of high school graduates in Equation 3 to estimate one's occupation in adulthood.

Therefore, the distribution of individuals from different social origins, conditional upon their education, into different social destinations is shaped by at least two educational processes, the association between children's education and occupation status in adulthood (captured by γ_g) and the educational distribution (α_g). In other words, the structures of the models suggest that, for the role of education-destination association shown by Figure 2, there are at least two sources – a). the "true changes" in education-destination association (γ_g) and b). some additional influences relevant to the educational upgrading (changed α_g). The descriptive pattern could not demonstrate the pure contribution of our process #2, education-destination association.

Within-Education-Group Gaps

Table 2 investigates the role of gaps within education groups (in social destination by social origin, process #3) in the third row. Different from the other educational processes, point estimates from the three ways of evaluation vary in both the direction and the magnitude of contribution, and none of the three evaluations yield a statistically significant contribution. Such disagreements may originate from the dissimilar changes among different groups, that the gaps narrowed for high school dropouts and graduates but widened for college graduates (Figure 3 and Appendix I – Table A4). Therefore, the contribution of within-education-group gaps to the cross-cohort change in intergenerational persistence seems still unclear.

Shares of Education Groups

Lastly, the estimated roles of changes in the shares of education groups are recorded in the fourth row of Table 2. Surprisingly, model simulations suggest that the increasing shares of college graduates, among whom there is lower persistence, in the post-expansion period *increased* intergenerational occupation persistence (with the contribution being statistically significant in the third way). This is different from theoretical expectations (Breen 2010; Hout 1988). However, compared to college graduates in the older cohort, the within-group persistence of college graduates in the younger cohort increased (Figure 3 and Appendix I – Table A4). These internal changes among college graduates ("AA" and "BA+") may have made the changing shares of education groups increase intergenerational occupation persistence, rather than decrease as expected.

Simulated Results in General

Putting the results in all four rows of Table 2 together, it is obvious that, at least in terms of the point estimates, the contributions of the four education processes offset each other. While the origin-education and education-destination associations tended to decrease intergenerational occupation persistence, the other two processes – within-education-group gaps and percentages of education groups – tended to increase persistence. On the other hand, the contributions of these processes are only statistically

significantly different from zero in the third way of evaluation that takes more pairs of comparisons into account (95% confidence intervals not overlapping zero). Therefore, in general, we don't enough evidence to reject the null hypothesis that none of the educational processes contributed to change in intergenerational persistence. However, this lack of statistical power does not contradict the point estimates but suggests that, when more data are available, further analyses are necessary to truly understand the multiple roles of education in intergenerational occupation persistence in postsocialist China.

Divergence in Descriptive and Model-Simulated Estimates

Considering these expected contributions of educational processes on the cross-cohort change in intergenerational occupation persistence from model simulations, there is one more question remaining for us – why the model-estimated contributions are divergent from the theory expectations and descriptive patterns? As roughly discussed above, the key issue is the educational upgrading (changes in shares of education groups, process #4, that higher percentages of the population are shifted to the groups of college graduates). I suspected that the educational upgrading made the descriptive results themselves not reflective of the underlying true and pure contributions of the other educational processes.

Model-Simulated Roles with and without Educational Upgrading

To investigate whether educational upgrading indeed adjusted the contributions of the other educational processes, I looked at all eight pairs of comparisons with which the model simulations estimate the roles of origin-education association (#1), education-destination association (#2), and within-education-group gaps (#3). Specifically, I compared the estimated contributions of these three educational processes if the educational upgrading took place as observed and the estimations if there was not the educational upgrading. To preview, I found that the roles of origin-education and education-destination associations, as well as gaps within education groups, are dissimilar with and without educational upgrading.

Figure 5 records the estimated contributions of educational processes under different counterfactual scenarios. For each panel, there are two variables of interest – the educational process listed above the panel and educational upgrading. The bars in each panel are arranged in a descending order. In addition, the darker-gray bars mark the scenarios with the educational upgrading, while the lighter-gray bars represent scenarios *without* educational upgrading. The statuses of the other two educational processes are held in control.

The two panels on the top of Figure 5 show model-simulated roles of origin-education and education-destination associations (process #1 and #2) when there was and was not the educational upgrading. The darker-gray bars are generally longer compared

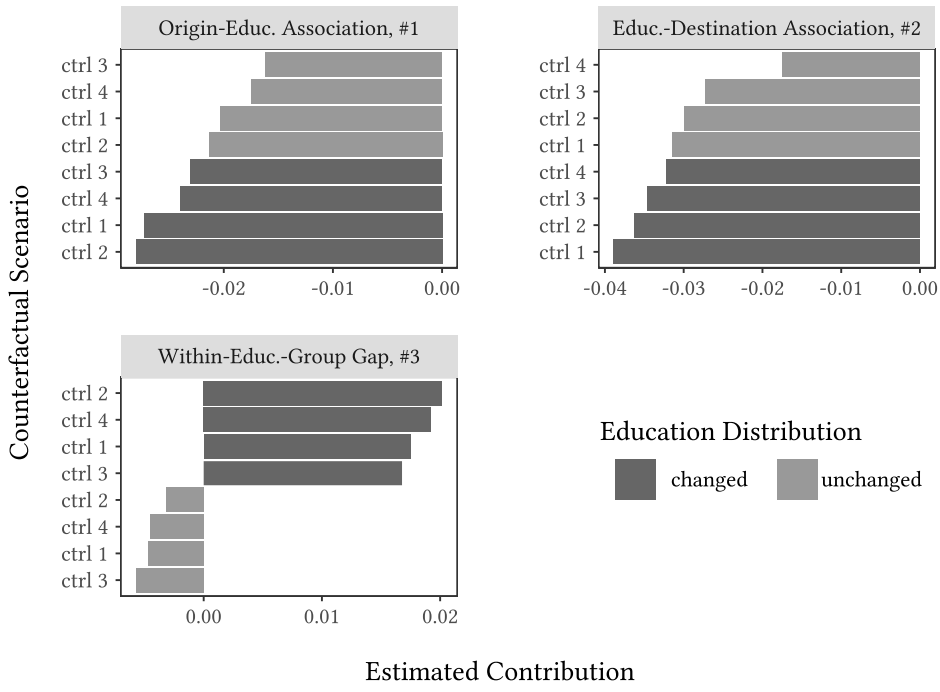


Figure 5. Model-Simulated Contributions of Educational Processes under Different Counterfactual Scenarios by Status of Educational Distribution

Source: CGSS 2008~2021 Data.

Note: Estimated values and details of counterfactual scenarios are in Appendix I – Table A9.

to the bars of lighter gray, indicating that the role of decreasing persistence tends to be stronger with the educational upgrading than *without* educational upgrading.

The bottom-left panel of Figure 5 shows model-simulated roles of within-education-group gaps (process #3). When the educational upgrading took place as observed, the models estimated the contribution to be increasing persistence; on the other hand, when educational distributions remained unchanged, model simulations suggested that the changed within-education-group gaps decrease persistence. Such countervailing results also explain why the three ways of estimations returned results that vary in both the direction and the magnitude of contribution (third row of Table 2).

Together, these three panels in Figure 5 have demonstrated that the model-simulated contributions of origin-education and education-destination associations (process #1 and #2), as well as within-education-group gaps (process #3), are dissimilar in scenarios where there is educational upgrading, compared to scenarios without the educational upgrading, in terms of magnitude and/or direction (increase or decrease). They qualitatively demonstrated that the educational upgrading that we observed likely adjusted the roles of other educational processes in the cross-cohort change of intergenerational occupation persistence, bringing some additional contributions. Therefore, since the descriptive patterns cannot exclude these additional influences following educational upgrading and reveal the true contributions of educational processes, we observed the divergence between descriptive and model-simulated results.

An Additional Test

In additional to the qualitative comparisons above, I also ran statistical tests to see if I can confirm that the educational upgrading adjusted the contributions of other educational processes. I again relied on the model parameters, but did not look at educational upgrading only. I fully tested whether the fact that one educational process evolved has adjusted the contribution of another educational process in the cross-cohort change of intergenerational occupation persistence. The results are recorded in Table 3. With a similar idea as the three ways of evaluation in Table 2, the first two ways treat the

factors to be controlled as changed and unchanged in the post-expansion period, respectively. The third way takes more information into account and considers all counterfactual scenarios. (For details of the approach, see Appendix III – The Model-Based Approach – Counterfactual Analyses.)

Table 3. Two-Way Interactions among Pairs of Educational Processes by Model Simulations

		As Only Unchanged	As Only Changed	Full Average of All
	Origin-Education Association	+.004	+.004	+.002
×	Education-Destination Association	(–.042, +.051)	(–.043, +.052)	(–.010, +.014)
	Origin-Education Association	–.001	–.001	–.001
×	Within-Education-Group Gaps	(–.049, +.048)	(–.049, +.047)	(–.012, +.011)
	Origin-Education Association	+.002	–.007	–.003
×	Shares of Education Group	(–.045, +.005)	(–.054, +.040)	(–.015, +.009)
	Education-Destination Association	+.002	+.001	+.001
×	Within-Education-Group Gaps	(–.045, +.050)	(–.046, +.048)	(–.011, +.013)
	Education-Destination Association	–.006	–.057	–.003
×	Shares of Education Group	(–.053, +.041)	(–.115, 0.000)	(–.015, +.008)
	Within-Education-Group Gaps	+.024	+.022	+.011
×	Shares of Education Group	(–.024, +.072)	(–.024, +.069)	(0.000, +.023)

Source: CGSS 2008~2021 Data.

Note: 95% intervals from repeated simulations in parentheses. Column 1 treats the target factors as unchanged (using old parameters), while other two changed as observed. Column 2 treats the target factor as changed (using young parameters). Column 3 takes all scenarios into account. See Appendix III – The Model-Based Approach – Counterfactual Analyses for details.

In Table 3, each row pays attention to one pair of the educational processes. For example, the first row looked at the origin-education and education-destination

associations (process #1 and #2), to see if the fact that the origin-education association evolved as observed has altered the role of the education-destination association or, in the other direction, if changed education-destination association has adjusted the role of origin-education association. According to the point estimates, most of such influences on each other were very weak (close to zero), except for the pairs that involve the shares of education groups (process #4, or educational upgrading). For example, according to the second way (Column 2), educational upgrading made origin-education and education-destination associations decrease persistence to a larger extent (.007 and .057, respectively). In addition, all three ways of estimation show that educational upgrading enlarged the increase in persistence from the changed gaps within education groups by at least .011.

Therefore, in additional to the qualitative patterns in the previous section (Figure 5), this additional test and its results in Table 3 has provided further suggestive evidence that the educational upgrading likely adjusted the contribution of other educational processes in the cross-cohort change of intergenerational occupation persistence. However, again, I do not have enough statistical evidence to make confirmations. This does not contradict the point estimates, though; it rather suggests that further efforts are needed to truly understand the multiple roles of education in intergenerational occupation persistence in postsocialist China with more data.

Robustness Check[26]: Taking Gender into Account

Since prior research has discovered gender differences in intergenerational occupation persistence in China (Huo and Golley 2022; Xie et al. 2022), I repeated the analyses above separately on gender-specific samples (Appendix I – Tables A1 and A10) and tested for gender differences (Appendix I – Tables A11 and A12). While it statistically significantly declined for women, overall intergenerational persistence remained stable for men. Furthermore, for both genders, most of the model-simulated contributions of educational processes are statistically insignificant. The point estimates suggest that stability in persistence across cohorts resulted from offsetting contributions between the educational processes, though. In addition, in terms of the differences in persistence, its cross-cohort change, and the contribution of educational processes, few values other than the older cohort's overall persistence statistically significantly differed between women and men.

In short, when analyzing gender-specific samples, I still encountered the limitation of lacking statistical evidence. However, it again does not contradict the point estimates, but for one more time suggests that further efforts are needed when new data are available to make more definitive claims. In addition, connecting these gender-specific patterns to that of the full sample, I again find that, at least in terms of the point estimates, the cross-cohort changes in persistence resulted from offsetting contributions between the educational processes.

Discussion

In the midst of rapid social, economic, and educational change in China in recent decades (Xie 2011; Xie 2012; Xie and Zhou 2014; Zheng, Zhang, and Zhu 2021; Zhou 2014; Zhou and Xie 2019), prior research has found little change in intergenerational persistence (Chen 2013; Xie et al. 2022; Zhou and Xie 2019). This finding is perplexing, given that prior research has also established rising educational stratification by social origins (Ding et al. 2021; Hu and Hibel 2014; Li 2014; Wu 2017; Wu 2013). In this thesis, I investigate how changing educational processes contributed to cross-cohort stability in intergenerational persistence in offsetting ways.

Using a model-based analytic approach from Bloome et al. (2018) to conduct my investigations, I decomposed the multifaceted roles of education in intergenerational occupation persistence and quantified the contribution of each ingredient to cross-cohort change in intergenerational persistence. My analyses relied on data from nine waves of Chinese General Social Survey.

Consistent with prior discoveries (Chen 2013; Xie et al. 2022; Zhou and Xie 2019), I found that intergenerational occupation persistence did not change substantially between cohorts who came of age prior to educational expansion (1978-1998) versus after expansion (1999-2021) in postsocialist China. Descriptive patterns suggested that offsetting contributions of the educational processes led to such a stability in persistence. In support of prior research (Ding et al. 2021; Hu and Hibel 2014; Hu and Hibel 2015; Wu

2017; Wu 2013), the origin-education and education-destination associations (process #1 and #2) strengthened, which were expected to increase intergenerational persistence. On the other hand, while the role of the gaps in social destination by social origin within education groups (process #3) was unclear, the educational upgrading (increased shares of college graduates, changes in process #4) was expected to decrease persistence according to the theory of Breen (2010:368).

I then proceeded to my main contribution – to formally estimate the contribution of each educational process with the model simulations. Although the results also suggested offsetting contributions, there were discrepancies compared to the descriptive estimates. The point estimates suggested that the origin-education and education-destination associations (process #1 and #2) decreased persistence. Indicating a reduced correlation, this role to weaken persistence of education-destination association supports the argument of Yang (2024) that the association between higher education credentials (educational attainment) and the high-status occupations (social destination) has diminished. In support of my hypothesis, changed within-education-group gaps in social destination by social origin (process #3) were estimated to increase persistence. In addition, educational upgrading (changes in the shares of education groups, process #4) was also predicted to increase persistence, which may led us to place a tiny question mark next to the argument of Breen (2010:328) that the growths of groups of higher education bring lower intergenerational persistence. I did not have enough statistical evidence to

confirm the individual contributions of each educational process or that offsetting contributions of educational processes led to stable persistence, though.

As a robustness check, I also applied the model simulations to gender-specific, *hukou*-specific, and gender-and-*hukou*-specific samples. Results also suggested that stability in intergenerational persistence likely resulted from the offsetting contributions of educational processes. However, again, I did not have enough statistical evidence to confirm the contributions of educational processes, identify the existence of differences by gender and/or *hukou* in the roles of education, or support this argument of offsetting contributions.

In addition to exploring the mechanism of China's social structure evolution through the lens of education's multifaceted roles in intergenerational occupation persistence, the results of my thesis can also be read in conversation with those of Bloome et al.'s (2018) analysis of the United States. In addition to lacking statistical power to confirm my findings, my main results differed from theoretical hypotheses and descriptive results, which were not observed in the case of the US. Why did we observe such a discrepancy? To put it in one sentence, the model simulations, but not the descriptive patterns, are capable of capturing the true contributions of educational processes.

Considering the uniqueness of China's social transformation, various social changes coincided and happened very rapidly (Xie 2011), therefore influencing the social impacts

of each other. Since the "Reform and Opening-up" initiated in 1978, it is only five decades; and in this short period of five decades (compared to the industrialization in the US started in late 18th century and lasted over a century), Chinese economic output grew tremendously. China has also completed the demographic transition from a high-fertility, high-mortality nation to a low-fertility, low-mortality country. More importantly, bringing the population structure with a small portion of people inactive in labor force, this demographic transition has benefited China's economic growth. The educational expansion was even more rapid – in only two decades (since 1999), the scale of higher education ten-folded (Zheng et al. 2021). By comparison, educational expansion took nearly a century in the United States (Fischer and Hout 2006). These pieces of evidence that Xie (2011) concluded suggest that coinciding social changes in China not only happened rapidly but also affected the impacts of each other – for example, the dramatic economic growth was due not only to economic evolutions (e.g., industrialization) but also to demographic transitions.

The hint for us on education's multiple roles in intergenerational occupation persistence trends is that changes in one educational process might have also adjusted the contribution of another. In my thesis, I compared the contributions of origin-education and education-destination associations (process #1 and #2), as well as within-education-group gaps (process #3), with and without educational upgrading (changed shares of education groups, process #4). These results demonstrated a qualitative pattern

where educational upgrading indeed altered the roles of the other three educational processes; in other words, if there was *not* the educational upgrading, we would receive different contributions in estimation from our current results. I also tested all pairs of educational processes to see if evolution in one of them adjusted the role of the other. The point estimates provided some support on such impacts on each other, though I was still unable to make confirmations.

Therefore, since the model simulations but not the descriptive patterns could distinguish the educational processes' true contributions and additional influences brought by others, we observed the discrepancy between model-simulated and descriptive results. In addition, given the potential influences of changed educational processes on the contribution in persistence trends of another, I suggest that future research further expand the four-component frame which Bloome et al. (2018) identified and on which I relied in my thesis, to better and more comprehensively unpack the multifaceted roles of education in the inter-generational persistence of socioeconomic (dis)advantages.

One limitation of my analyses is that the results were often not statistically significant. I could not confirm the contributions of the four educational processes to the cross-cohort change in intergenerational occupation persistence of if their contributions offset each other, nor could I test if changes in one educational process may adjust the

contributions of another. When new data are available, the larger sample sizes for analyses will increase statistical power and allow scholars to make more definitive claims.

In addition, future research should consider additional sources of heterogeneity in changes in intergenerational persistence. Though I only briefly reported the results with women and men tested separately in the text, I also ran the analyses for sub-samples defined by *hukou* status (household registration), and the results are in Appendix I. In addition to gender and *hukou*, scholars may also want to consider regional disparities in future research, given that the experiences of people in eastern, urban, coastal regions are likely different from those in western, rural, inland places (Asadullah and Xiao 2020; Qin et al. 2020; Shi et al. 2024).

More broadly, my thesis has not established causal relationships. Research of intergenerational occupation persistence concerns the extent to which inequalities and (dis)advantages in parental occupation status (social origin) are reproduced in adulthood, rather than the causal effect on parental occupation on adult occupation. Therefore, I (as well as the methodological framework of Bloome et al. 2018) followed in this descriptive tradition in intergenerational studies and document how sociodemographic changes in the population are associated with cross-cohort changes in persistence. My analyses show that changes in the association between parents' occupation status and children's educational attainment, the correlation between children's education and occupation status in adulthood, group persistence reflected by gaps in adult occupation status by

parental occupation status among people with the same education, and shares of college graduates in the population likely contributed to changes in intergenerational occupation persistence. Therefore, in addition to confirming the explicit roles of educational processes, future research in the context of China may also want to tease out to what extent these relationships are causal. With these further efforts, we may discover whether the descriptive patterns I have presented in this thesis have over- or under-stated the tendency of education to disrupt or reproduce the transmission of family (dis)advantages across generations in contemporary China and, therefore, deepen our understanding of China's social changes and the mechanism of social structure evolutions.

FOOTNOTES

1. For clarity, I use "persistence" throughout this thesis. I discuss persistence because it directly corresponds to the measurement of interest, which is the rank-rank linear regression slope between parents' and children's occupation statuses. Mobility is the complement of this slope, and, mathematically, $\text{mobility} = 1 - \text{persistence}$.
2. *Gaokao*, as the major channel of admission to higher education institutions in contemporary China, was initiated in 1952. It was temporarily cancelled 1966~1976. While the first resumed *gaokao* in 1977 was conducted in December, the country has been organizing the annual examinations in the summer (June or July) since 1978.
3. *Guanxi* can also be described as the differential model of association in traditional Chinese society, a structure that distinguishes interpersonal relationships based on the degree of closeness and distance (Fei 1992). Kinship relationships, including blood and marriage, are at the core of *guanxi*; in the periphery of kinship are non-kinship relationships such as classmates, neighbors, and fellow villagers. Kinship is more important than non-kinship; the farther away a relationship is from kinship, the less important it is. Meanwhile non-kinship ties can be upgraded to stable and intimate *guanxi* as the involved parties increase their favors and obligations. Important moments and scenarios, such as traditional and statutory holidays,

weddings, birthday parties, and social eating, are all important opportunities for constructing and maintaining *guanxi* (Bian 2001). For more details, see Yang and Zhang (2024:3-4, Section 2.2 and Table 1); Cheng (2024:327-328); Bian and Yang (2019); and Zhai (2009).

4. In the statistical bulletins (written in Chinese), the number of enrolled undergraduate students is the number of "本专科在校生," while that of enrolled graduates is the number of "在学研究生." The "total population" is the population number at the end of the year ("年末全国总人口").

5. This consists of "Project 985" universities, "Project 211" universities that are not "Project 985" institutions, 4-year institutions founded before 1999, 4-year institutions founded after 1999, and 3-year vocational institutions. The "Project 985" and "Project 211" institutions are generally the elite universities (Li 2004:17). The "Project 985" was named because it was initiated after a speech in May 1998, and its synonym is "The Project for Founding World-Class Universities." There are 39 universities in this category (Department of Degree Management 2008b). The "Project 211," starting in 1995, was named because it aimed at strengthening about 100 universities as a national priority for the 21st century. There are 112 universities in total as Project-211 universities, including 39 Project-985 universities and other 73 universities that are not Project-985 universities (Department of Degree Management 2008a). In short, in the post-expansion period,

among all elite universities, Project-985 universities are generally better than non-Project-985 Project-211 universities; the 4-year institutions, founded before and after 1999, are "ordinary" institutions in the middle; and the 3-year vocational institutions, most of whose students graduate with degrees equivalent to associate degrees, are at the very bottom.

6. In the 1998 statistical bulletins (written in Chinese), the number of enrolled undergraduate students is the number of "本专科在校生," while that of enrolled graduates is the number of "在学研究生." The "total population" is the population number at the end of the year ("年末全国总人口"). In 2022, the total enrollment is termed "各种形式的高等教育在学总规模," and the population is "年末全国人口."
7. I count "毕业研究生" and "普通本科毕业生" for 2022.
8. Since both of my cohorts consist of a span of 16 years, which represent wider age ranges than those in Bloome et al. (2018), I need to control for age. Therefore, technically, I need two linear regressions for the estimations of intergenerational occupation persistence. First, I predict the adult occupational ranks with both age and the square of age (age^2). Next, I calculate the residuals in the first regression and, in another linear regression, predict these residuals with parental ranks. The slope coefficient in this second regression captures intergenerational occupation persistence. However, the estimations with these two regressions and those with Equation 1 are almost the same (Appendix I – Table A1). Consequently, for

simplicity, I just use one linear regression (Equation 1) to estimate persistence. I do not control for age in the models, either.

9. I re-simulated for 21,120 times, and take the 2.5th and the 97.5th percentiles of the repeated simulation distribution to estimate the 95% intervals.

10. The statuses of β_g and v_g change together. They are in one group of parameters to demonstrate the status of one variable – gaps within education groups.

11. There are three educational processes to control, and each of these processes has two statuses. Therefore, the number of random combinations is $2^3 = 8$.

12. In this scenario, the counterfactual outcome is simulated by applying all four sets of parameters estimated from the older cohort on the younger cohort. It is not the older cohorts' simulated persistence, which is simulated by applying the older cohort's parameters on the older cohort.

13. CGSS collects both the respondents' current highest level of education and the status of that attainment (graduated or not). If the respondents have graduated, the answered highest level is his or her most up-to-date completed highest level of education at the time of the survey. If not, the actual highest level will be one level lower than the answer. For example, the actual highest level is "less than high school" when a respondent reported "high school" but the status was not "graduated." When a respondent reported "associate degree" or "bachelor's degree or above" but had not graduated, their actual highest level is "high school."

14. I need to ensure that the respondents were active in the labor force when they took the survey. According to Zhou and Xie (2019:1821), respondents needed to be 64 or younger. With the potentially oldest respondents born in 1964, the maximum age of individuals in my sample is $57 = 2021 - 1964$. Therefore, it is certain that all individuals in my sample were active in the labor force when they took the survey.
15. To ensure that all my respondents were 25 years old or older at the time of the survey, since the newest wave of CGSS was conducted in 2021, the potentially youngest respondents should be born in 1995.
16. A person's social origin is defined by the parental occupation when he or she was 14 (Zhou and Xie 2019:1821). Since none of my respondents were older than 14 in 1978, their social origins were all influenced by postsocialist China.
17. I convert ISCO-88 codes to ISEI scores with an "isco88conversion" package in R. It can be found at: https://rdrr.io/github/pdparker/isco88conversion/
18. In CGSS 2017, 2018, and 2021, the jobs are recorded with ISCO-08 codes. So, before determining their ISEI scores, I converted them back to ISCO-88 codes using a table from the International Labor Organization: https://ilostat.ilo.org/methods/concepts-and-definitions/classification-occupation/
19. There are also a number of agricultural workers whose ISCO codes are not provided by the raw CGSS data. I manually assign them a 6100, which is the most common value in the CGSS datasets if there is a code assigned. Though there are

some other values as well, like 6110, 6111, and 6114, the results remain accurate, because all these ISCI-88 codes are associated with the ISEI score 23 (Ganzeboom and Treiman 1996:221-238).

20. For example, in a group (e.g., the parents of older cohort individuals), there are 100 individuals in total. Among these 100 people, there are four distinct ISEI scores, from lowest to highest, *A*, *B*, *C*, and *D*. There are 40 people whose jobs are associated with ISEI *A*. 30, 20, and 10 individuals are matched with ISEI scores *B*, *C*, and *D*, respectively. After the three steps above, those with an ISEI *A* will be assigned the 20th percentile, the mid-point of 0 and 40. Those with ISEI *B* will be given the 55th percentile, the mid-point of 40 and 70. Similarly, those with ISEI *C* and *D* will be recorded as 80th percentile and 95th percentile, respectively.

21. A junior college program, or *zhuanke*, usually refers to a 2- or 3-year undergraduate program. A *zhuanke* degree is a postsecondary degree that is less valuable than a bachelor's degree in China, so it is similar to an associate's degree in the US higher education system. Thus, for simplicity, I borrow the concept and use "AA" as its abbreviation.

22. A bachelor's degree is awarded after a student completes a 4- or 5-year *benke* course.

23. The "BA+" category includes bachelor's degree, master's degree, doctoral degree, etc. Because of the small numbers of advanced degree (master's degree or higher,

yanjiusheng) holders in both cohorts (196, or 0.95%, in the older cohort; 292, or 3.45%, in the younger cohort), they are combined with BA holders.

24. Since the percentile range of jobs associated with ISEI 23 is roughly 8.66%~59.99% for the parents in the 1964~1979 cohort, it is difficult to reasonably divide people into the first and the second quartiles. Thus, I combine the lowest two quartiles as "Bottom Half." In other words, children from the "Bottom Half" category come from the bottom half of the parental occupation distribution.

25. Wald tests confirm that the correlations of occupation ranks in both cohorts vary statistically significantly with education.

26. I have analyzed gender-specific (women and men), *hukou*-specific (rural *hukou* holders and urban *hukou* holders), and gender-and-*hukou*-specific (women with rural *hukou*, women with urban *hukou*, men with rural *hukou*, and men with urban *hukou*) subsamples. Additional Tables are in Appendix I. To make cross-group comparisons, I do not re-rank them when extracting these sub-samples. In other words, I only rank the full sample once, and them separate this ranked full sample by gender and/or *hukou*.

APPENDIX I
Additional Tables and Figures

Table A1. Estimated Intergenerational Occupation Persistence by Cohort and its Cross-Cohort Change in Observation, Estimated with Age Controlled, and by Model Simulations

	Older Cohort (1964~1979)	Younger Cohort (1980~1995)	Δ $\beta^{young} - \beta^{old}$
Penal A: Point Estimates			
Full Sample			
Observed	.335	.310	–.025
Age Controlled	.328	.308	–.020
Model-Simulated	.326	.300	–.026
Women			
Observed	.371	.305	–.067
Age Controlled	.365	.301	–.065
Model-Simulated	.365	.308	–.057
Men			
Observed	.303	.315	+.013
Age Controlled	.295	.314	+.019
Model-Simulated	.292	.295	+.003
Rural Hukou Holders			
Observed	.206	.213	+.007
Age Controlled	.206	.213	+.007
Model-Simulated	.213	.208	–.005
Urban Hukou Holders			
Observed	.183	.239	+.056
Age Controlled	.176	.239	+.063
Model-Simulated	.225	.250	+.025
Women, Rural Hukou			
Observed	.217	.224	+.008
Age Controlled	.215	.222	+.007
Model-Simulated	.227	.225	–.001
Women, Urban Hukou			
Observed	.196	.205	+.008
Age Controlled	.193	.197	+.005
Model-Simulated	.258	.230	–.029
Men, Rural Hukou			
Observed	.197	.204	+.008

	Older Cohort (1964~1979)	Younger Cohort (1980~1995)	Δ $\beta^{young} - \beta^{old}$
Age Controlled	.196	.204	+.008
Model-Simulated	.200	.198	–.003
Men, Urban Hukou			
Observed	.166	.266	+.100
Age Controlled	.161	.269	+.108
Model-Simulated	.193	.265	+.072
Penal B: 95% Intervals			
Full Sample			
Observed	(.320, .350)	(.286, .334)	(–.053, +.004)
Age Controlled	(.313, .343)	(.284, .332)	(–.049, +.008)
Model-Simulated	(.311, .341)	(.276, .323)	(–.055, +.002)
Women			
Observed	(.349, .393)	(.272, .337)	(–.106, –.027)
Age Controlled	(.343, .387)	(.267, .333)	(–.105, –.026)
Model-Simulated	(.343, .386)	(.272, .342)	(–.098, –.016)
Men			
Observed	(.281, .324)	(.280, .349)	(–.029, +.053)
Age Controlled	(.273, .316)	(.279, .348)	(–.022, +.059)
Model-Simulated	(.271, .314)	(.262, .328)	(–.037, +.042)
Rural Hukou Holders			
Observed	(.186, .227)	(.180, .247)	(–.032, +.046)
Age Controlled	(.186, .226)	(.179, .246)	(–.031, +.046)
Model-Simulated	(.193, .233)	(.174, .241)	(–.045, +.033)
Urban Hukou Holders			
Observed	(.151, .214)	(.187, .294)	(–.004, +.119)
Age Controlled	(.145, .208)	(.188, .291)	(+.003, +.122)
Model-Simulated	(.187, .262)	(.193, .303)	(–.043, +.091)
Women, Rural Hukou			
Observed	(.188, .245)	(.180, .269)	(–.045, +.062)
Age Controlled	(.187, .243)	(.178, .267)	(–.045, +.061)
Model-Simulated	(.198, .255)	(.180, .272)	(–.055, +.053)
Women, Urban Hukou			
Observed	(.147, .247)	(.119, .304)	(–.092, +.118)
Age Controlled	(.145, .242)	(.114, .276)	(–.093, +.097)
Model-Simulated	(.203, .313)	(.130, .317)	(–.141, +.076)
Men, Rural Hukou			
Observed	(.168, .225)	(.155, .252)	(–.049, +.064)

	Older Cohort (1964~1979)	Younger Cohort (1980~1995)	Δ $\beta^{young} - \beta^{old}$
Age Controlled	(.168, .224)	(.155, .252)	(–.048, +.064)
Model-Simulated	(.172, .229)	(.150, .245)	(–.058, +.052)
Men, Urban Hukou			
Observed	(.125, .208)	(.202, .331)	(+.024, +.178)
Age Controlled	(.120, .202)	(.204, .333)	(+.032, +.184)
Model-Simulated	(.142, .244)	(.199, .328)	(–.010, +.153)

Source: CGSS 2008~2021 Data.

Note: Bootstrapped 95% intervals for observed and with-age-controlled values, and 95% intervals from repeated simulations for model-simulated values, in parentheses. The values with "age controlled" are calculated as described in Footnote #8 and Appendix III – Overall Persistence without and with Age Controlled.

Table A2. Estimated Values: Percentages of Individuals with Bachelor's or Advanced Degrees across Parental Occupational Quartiles and Cohorts

	% Bottom Half	% 3rd Quartile	% 4th Quartile
Older Cohort (1964~1979 group)	4.213	9.942	23.070
	(3.819, 4.627)	(8.938, 10.980)	(21.540, 24.621)
Younger Cohort (1980~1995 group)	14.566	24.940	48.394
	(13.004, 15.730)	(22.652, 27.179)	(45.821, 50.995)

Source: CGSS 2008~2021 Data.

Note: Bootstrapped 95% intervals in parentheses. The 1st and 2nd Quartiles of parents' occupational percentile ranks are combined as the "Bottom Half" category.

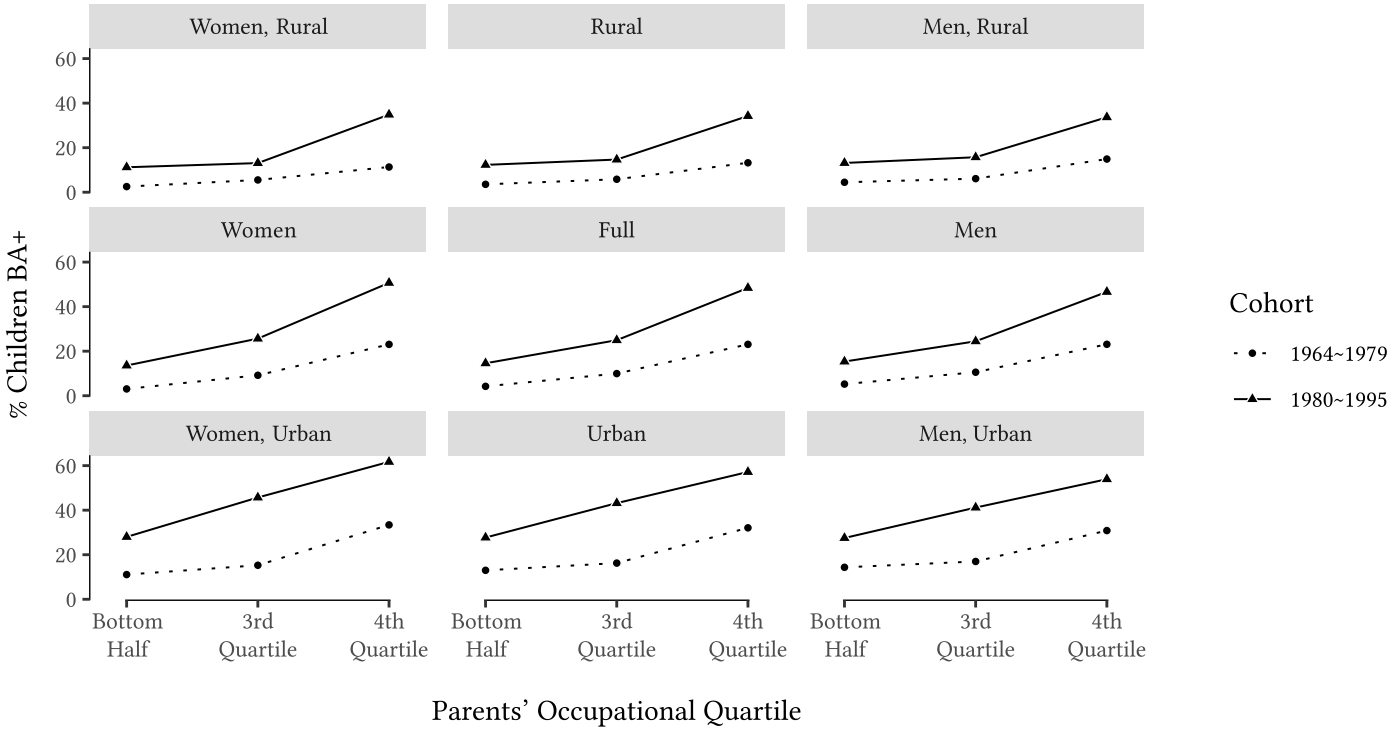


Figure A1. Percentages of Children with Bachelor's or Advanced Degrees across Parental Occupational Quartiles and Cohorts in Different Sub-Samples: the full sample (center), all women (middle left), all men (middle right), all rural *hukou* holders (top middle), all urban *hukou* holders (bottom middle), women with rural *hukou* (top left), women with urban *hukou* (bottom left), men with rural *hukou* (top right), and men with urban *hukou* (bottom right)

Source: CGSS 2008~2021 Data.

Note: The 1st and 2nd Quartiles of Quartiles of parents' occupation percentile ranks are combined as the "Bottom Half" category. The center panel (of full sample) is exactly the same as Figure 1, whose estimated values are in Appendix I – Table A2 above. Gender is on the horizontal axis, with women on the left and men on the right. *Hukou* status is on the vertical axis, with rural *hukou* holders on the top and urban *hukou* holders on the bottom.

Table A3. Estimated Values: Average Adult Occupation Rank Relative to High School Dropouts by Educational Attainment and Cohort

	<HS	HS	AA	BA+
	Relative Rank	Relative Rank	Relative Rank	Relative Rank
Older Cohort (1964~1979 group)	1	1.633	2.010	2.162
	(1, 1)	(1.604, 1.662)	(1.979, 2.040)	(2.137, 2.187)
Younger Cohort (1980~1995 group)	1	1.610	2.010	2.357
	(1, 1)	(1.546, 1.675)	(1.940, 2.083)	(2.294, 2.422)

Source: CGSS 2008~2021 Data.

Note: Bootstrapped 95% intervals in parentheses. High school dropouts (<HS) are individuals without a high school diploma, and their relative rank (to themselves) is set to 1.

Table A4. Group-Specific Regression Slopes by Educational Attainment and Cohort (Age Group) and their Cross-Cohort Changes

	Group-Specific Regression Slopes		95% Intervals	
	Point Estimates	Δ	Point Estimates	Δ
Older Cohort (1964~1979 group)				
<HS	.170		(.151, .189)	
HS	.116		(.080, .152)	
AA	.014		(–.029, .055)	
BA+	.017		(–.011, .046)	
Younger Cohort (1980~1995 group)				
<HS	.136	–.034	(.070, .125)	(–.076, –.009)
HS	.102	–.014	(.049, .154)	(–.079, +.049)
AA	.029	+.015	(–.028, .087)	(–.056, +.087)
BA+	.073	+.056	(.033, .113)	(+.007, +.105)

Source: CGSS 2008~2021 Data.

Note: Bootstrapped 95% intervals in parentheses.

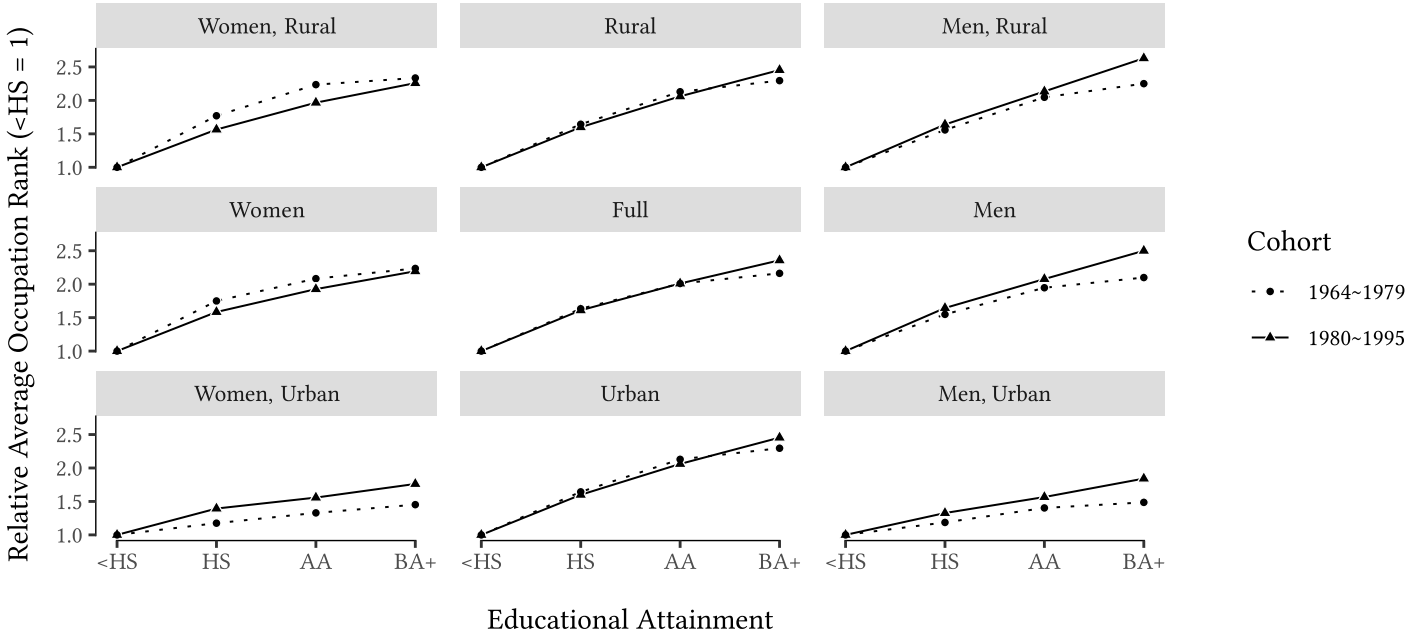


Figure A2. Average Adult Occupation Rank Relative to High School Dropouts by Educational Attainment and Cohort in Different Sub-Samples: the full sample (center), all women (middle left), all men (middle right), all rural *hukou* holders (top middle), all urban *hukou* holders (bottom middle), women with rural *hukou* (top left), women with urban *hukou* (bottom left), men with rural *hukou* (top right), and men with urban *hukou* (bottom right)

Source: CGSS 2008~2021 Data.

Note: High school dropouts (<HS) are individuals without a high school diploma, and their relative rank (to themselves) is set to 1. The center panel (of full sample) is exactly the same as Figure 2, whose estimated values are in Appendix I – Table A3 above. Gender is on the horizontal axis, with women on the left and men on the right. *Hukou* status is on the vertical axis, with rural *hukou* holders on the top and urban *hukou* holders on the bottom.

Table A5. Group Shares by Educational Attainment and Cohort (Age Group) and their Cross-Cohort Changes

	Shares %		95% Intervals	
	Point Estimates	Δ	Point Estimates	Δ
Older Cohort (1964~1979 group)				
<HS	68.8		(68.1, 69.5)	
HS	14.1		(13.6, 14.6)	
AA	8.2		(7.7, 8.6)	
BA+	8.9		(8.4, 9.3)	
Younger Cohort (1980~1995 group)				
<HS	40.7	−28.2	(39.4, 42.0)	(−29.6, −26.7)
HS	18.7	+4.6	(17.7, 19.7)	(+3.5, +5.7)
AA	15.7	+7.5	(14.8, 16.6)	(+6.5, +8.5)
BA+	24.9	+16.1	(23.8, 26.0)	(+14.9, +17.2)

Source: CGSS 2008~2021 Data.

Note: Bootstrapped 95% intervals in parentheses.

Table A6. Coefficients of Generalized Ordered Logistic Regression Predicting Educational Transitions with Parental Occupation Rank

	Older Cohort (1964~1979)		Younger Cohort (1980~1995)	
	α	λ	α	λ
	Share Parameter	Orig.-Edu. Parameter	Share Parameter	Orig.-Edu. Parameter
Full Sample				
≥ HS	−0.891	2.853	+0.428	2.561
	(−0.930, −0.852)	(2.712, 2.996)	(+0.368, +0.488)	(2.352, 2.778)
≥ AA	−1.786	2.976	−0.423	2.586
	(−1.836, −1.736)	(2.809, 3.144)	(−0.483, −0.364)	(2.385, 2.793)
≥ BA+	−2.579	2.900	−1.228	2.561
	(−2.648, −2.512)	(2.681, 3.121)	(−1.298, −1.161)	(2.335, 2.793)

	Older Cohort (1964~1979)		Younger Cohort (1980~1995)	
	α Share Parameter	λ Orig.-Edu. Parameter	α Share Parameter	λ Orig.-Edu. Parameter
Women				
≥ HS	−1.121	3.394	+0.428	2.790
	(−1.182, −1.061)	(3.184, 3.609)	(+0.339, +0.517)	(2.483, 3.108)
≥ AA	−1.954	3.429	−0.356	2.849
	(−2.035, −1.876)	(3.179, 3.688)	(−0.445, −0.269)	(2.557, 3.157)
≥ BA+	−2.803	3.488	−1.249	2.874
	(−2.919, −2.696)	(3.162, 3.839)	(−1.352, −1.150)	(2.546, 3.217)
Men				
≥ HS	−0.720	2.467	+0.428	2.395
	(−0.772, −0.671)	(2.279, 2.654)	(+0.348, +0.511)	(2.118, 2.682)
≥ AA	−1.662	2.645	−0.473	2.400
	(−1.729, −1.598)	(2.426, 2.868)	(−0.555, −0.395)	(2.132, 2.681)
≥ BA+	−2.425	2.496	−1.215	2.334
	(−2.514, −2.342)	(2.209, 2.786)	(−1.310, −1.127)	(2.036, 2.644)
Rural *Hukou* Holders				
≥ HS	−1.443	1.672	−0.134	1.488
	(−1.491, −1.397)	(1.488, 1.855)	(−0.198, −0.070)	(1.236, 1.745)
≥ AA	−2.323	1.899	−0.976	1.494
	(−2.388, −2.260)	(1.673, 2.124)	(−1.049, −0.904)	(1.229, 1.765)
≥ BA+	−3.048	1.890	−1.746	1.609
	(−3.140, −2.960)	(1.573, 2.201)	(−1.841, −1.658)	(1.284, 1.939)
Urban *Hukou* Holders				
≥ HS	+0.910	2.061	+2.115	2.620
	(+0.837, +0.983)	(1.773, 2.354)	(+1.949, +2.295)	(1.972, 3.307)
≥ AA	−0.319	2.054	+0.860	2.506
	(−0.390, −0.249)	(1.738, 2.387)	(+0.748, +0.971)	(2.083, 2.959)
≥ BA+	−1.316	2.129	−0.191	2.234
	(−1.404, −1.233)	(1.742, 2.557)	(−0.295, −0.093)	(1.835, 2.666)
Women, Rural *Hukou*				
≥ HS	−1.767	+1.957	−0.150	1.799
	(−1.841, −1.694)	(1.682, 2.234)	(−0.242, −0.058)	(1.447, 2.168)
≥ AA	−2.574	2.066	−0.948	1.769
	(−2.680, −2.475)	(1.727, 2.401)	(−1.055, −0.846)	(1.388, 2.163)
≥ BA+	−3.342	2.318	−1.828	1.872
	(−3.497, −3.204)	(1.832, 2.795)	(−1.970, −1.699)	(1.374, 2.369)

	Older Cohort (1964~1979)		Younger Cohort (1980~1995)	
	α	λ	α	λ
	Share Parameter	Orig.-Edu. Parameter	Share Parameter	Orig.-Edu. Parameter
Women, Urban *Hukou*				
≥ HS	+1.027	2.664	+2.218	2.650
	(+0.919, +1.140)	(2.234, 3.188)	(+1.920, +2.540)	(1.316, 3.948)
≥ AA	−0.239	2.419	+1.097	2.737
	(−0.342, −0.135)	(1.959, 2.921)	(+0.899, +1.292)	(2.022, 3.514)
≥ BA+	−1.312	2.533	−0.074	2.489
	(−1.448, −1.190)	(1.952, 3.208)	(−0.245, +0.084)	(1.871, 3.159)
Men, Rural *Hukou*				
≥ HS	−1.205	1.499	−0.122	1.253
	(−1.266, −1.146)	(1.255, 1.742)	(−0.211, −0.033)	(0.907, 1.607)
≥ AA	−2.136	1.796	−0.998	1.281
	(−2.222, −2.056)	(1.500, 2.093)	(−1.101, −0.899)	(0.915, 1.654)
≥ BA+	−2.844	1.631	−1.687	1.419
	(−2.964, −2.736)	(1.212, 2.033)	(−1.817, −1.569)	(0.992, 1.846)
Men, Urban *Hukou*				
≥ HS	+0.823	1.602	+2.405	2.615
	(+0.725, +0.924)	(1.212, 1.994)	(+1.857, +2.263)	(1.976, 3.299)
≥ AA	−0.385	1.762	+0.703	2.412
	(−0.482, −0.289)	(1.335, 2.208)	(+0.569, +0.839)	(1.895, 2.968)
≥ BA+	−1.324	1.841	−0.277	2.065
	(−1.445, −1.212)	(1.320, 2.402)	(−0.409, −0.151)	(1.541, 2.629)

Source: CGSS 2008~2021 Data.

Note: Bootstrapped 95% intervals in parentheses.

Table A7. Coefficients and Residual Standard Deviations of Linear Regressions Predicting Adult Occupation ISEI Scores with Parental Occupation Rank

	Older Cohort (1964~1979)			Younger Cohort (1980~1995)		
	γ Edu.-Dest. Parameter	β Slope Parameter	υ Residual SD Parameter	γ Edu.-Dest. Parameter	β Slope Parameter	υ Residual SD Parameter
Full Sample						
<HS	30.094	7.310***	10.238	33.642	6.386***	10.635
	(29.879, 30.311)	(+6.421, +8.217)	(10.040, 10.434)	(33.153, 34.140)	(+4.560, +8.257)	(10.307, 10.966)
HS	41.510	5.568***	15.462	42.242	5.250***	13.432
	(40.883, 42.153)	(+3.338, +7.779)	(15.057, 15.863)	(41.489, 43.006)	(+2.452, +8.104)	(12.906, 13.944)
AA	51.116	−0.434	13.722	48.575	0.973	13.396
	(50.199, 52.027)	(−3.734, +2.707)	(13.247, 14.182)	(47.664, 49.489)	(−2.345, +5.264)	(12.839, 13.917)
BA+	57.319	−0.455	14.677	54.609	3.665***	13.971
	(56.338, 58.295)	(−3.345, +2.476)	(14.245, 15.088)	(53.828, 55.396)	(+1.162, +6.243)	(13.576, 14.351)
Women						
<HS	29.833	7.688***	10.333	34.712	6.407***	11.046
	(29.543, 30.132)	(+6.444, +8.953)	(10.076, 10.586)	(33.971, 35.455)	(+3.771, +9.039)	(10.622, 11.482)
HS	42.656	5.595***	14.590	43.279	1.313	12.422
	(41.731, 43.581)	(+2.471, +8.708)	(13.998, 15.165)	(50.163, 52.880)	(−5.464, +5.303)	(12.246, 13.620)
AA	51.515	0.083	12.942	48.926	2.023	11.848
	(50.163, 52.880)	(−5.464, +5.303)	(12.246, 13.620)	(47.840, 50.015)	(−1.710, +5.777)	(11.106, 12.558)
BA+	57.114	2.492	14.211	54.146	4.056**	13.211
	(55.654, 58.623)	(−1.698, +6.662)	(13.568, 14.819)	(53.097, 55.189)	(+0.756, +7.349)	(12.661, 13.738)
Men						

	Older Cohort (1964~1979)			Younger Cohort (1980~1995)		
	γ	β	υ	γ	β	υ
	Edu.-Dest. Parameter	Slope Parameter	Residual SD Parameter	Edu.-Dest. Parameter	Slope Parameter	Residual SD Parameter
<HS	30.344	6.920***	10.130	32.843	6.614***	10.186
	(30.034, 30.659)	(+5.642, +8.225)	(9.829, 10.435)	(32.199, 33.500)	(+4.167, +9.903)	(9.682, 10.686)
HS	40.793	4.996***	16.061	41.549	7.743***	14.068
	(39.944, 41.646)	(+1.942, +8.010)	(15.494, 16.620)	(40.523, 42.578)	(+3.899, +11.648)	(13.337, 15.547)
AA	50.809	−1.005	14.303	48.260	0.179	14.801
	(49.497, 52.025)	(−4.851, +2.803)	(13.648, 14.916)	(46.831, 49.684)	(−4.911, +5.276)	(13.992, 15.547)
BA+	57.318	−2.517	14.985	54.946	3.449*	14.629
	(56.021, 58.590)	(−6.310, +1.397)	(14.393, 15.548)	(53.823, 56.090)	(−0.228, +7.125)	(14.065, 15.169)
Rural *Hukou* Holders						
<HS	29.007	5.281***	9.719	32.739	5.354***	10.295
	(28.803, 29.214)	(+4.335, +6.268)	(9.507, 9.930)	(32.273, 33.220)	(+3.418, +7.328)	(9.969, 10.629)
HS	40.959	4.862***	16.436	41.242	3.802*	13.375
	(40.100, 41.843)	(+1.683, +8.056)	(15.861, 17.005)	(40.316, 42.182)	(+0.140, +7.585)	(12.713, 14.027)
AA	52.136	3.111	14.739	48.504	−0.438	14.122
	(50.889, 53.348)	(−1.144, +7.376)	(14.025, 15.411)	(47.205, 49.806)	(−5.233, +4.329)	(13.325, 14.852)
BA+	58.786	−1.024	14.866	55.498	4.114*	14.758
	(57.478, 60.088)	(−5.025, +3.077)	(14.184, 15.513)	(54.350, 56.664)	(+0.442, +7.771)	(14.154, 15.343)
Urban *Hukou* Holders						
<HS	37.210	5.087***	12.263	38.351	6.840**	12.513
	(36.475, 37.981)	(+2.099, +8.094)	(11.711, 12.801)	(36.744, 40.027)	(−0.158, +13.255)	(11.277, 13.653)
HS	42.980	5.580***	14.126	44.561	5.482**	13.351
	(42.148, 43.806)	(+2.299, +8.887)	(13.573, 14.662)	(43.294, 45.838)	(+0.961, +10.212)	(12.498, 14.130)

	Older Cohort (1964~1979)			Younger Cohort (1980~1995)		
	γ Edu.-Dest. Parameter	β Slope Parameter	υ Residual SD Parameter	γ Edu.-Dest. Parameter	β Slope Parameter	υ Residual SD Parameter
AA	49.924	−0.366	12.700	48.805	2.842	12.626
	(48.810, 51.053)	(−5.990, +4.994)	(12.050, 13.325)	(47.629, 50.001)	(−2.483, +7.958)	(11.831, 13.361)
BA+	55.622	6.018**	14.427	54.055	7.275***	13.269
	(54.498, 56.726)	(+0.917, +11.319)	(13.839, 14.972)	(53.130, 54.975)	(+3.197, +11.433)	(12.749, 13.768)
Women, Rural *Hukou*						
<HS	28.659	5.528***	9.707	33.904	5.281***	10.899
	(28.386, 28.935)	(+4.247, +6.845)	(9.432, 9.980)	(33.197, 34.614)	(+2.493, +8.056)	(10.456, 11.355)
HS	42.102	5.308*	15.612	42.256	0.899	11.945
	(40.837, 43.376)	(+0.824, +9.777)	(14.767, 16.421)	(41.018, 43.493)	(−3.356, +5.157)	(11.075, 12.772)
AA	52.858	6.155*	13.670	48.689	2.800	12.063
	(50.927, 54.789)	(+0.110, +12.272)	(12.575, 14.672)	(47.172, 50.175)	(−2.519, +8.109)	(11.051, 13.004)
BA+	57.739	−0.662	14.807	55.172	4.052	14.167
	(55.722, 59.754)	(−6.876, +5.513)	(13.742, 15.787)	(53.589, 56.758)	(−1.003, +9.032)	(13.279, 14.968)
Women, Urban *Hukou*						
<HS	38.759	5.766**	12.761	39.383	9.866**	11.844
	(37.609, 39.928)	(+0.810, +10.615)	(12.026, 13.464)	(36.903, 42.123)	(−2.994, +19.143)	(10.246, 13.249)
HS	44.124	5.120*	13.449	45.639	−1.808	12.972
	(42.952, 45.298)	(+0.350, +10.069)	(12.622, 14.241)	(43.550, 47.783)	(−8.130, +4.711)	(11.621, 14.214)
AA	50.272	0.005	12.280	49.314	0.733	11.644
	(48.557, 51.914)	(−8.679, +8.307)	(11.359, 13.175)	(47.878, 50.812)	(−5.596, +6.666)	(10.526, 12.684)
BA+	56.971	7.927*	13.777	53.598	8.067***	12.434
	(55.432, 58.503)	(+0.737, +15.156)	(12.952, 14.538)	(52.481, 54.737)	(+3.419, +12.645)	(11.729, 13.112)

	Older Cohort (1964~1979)			Younger Cohort (1980~1995)		
	γ Edu.-Dest. Parameter	β Slope Parameter	υ Residual SD Parameter	γ Edu.-Dest. Parameter	β Slope Parameter	υ Residual SD Parameter
Men, Rural *Hukou*						
<HS	29.348	5.030***	9.724	31.835	5.645***	9.600
	(29.047, 29.656)	(+3.614, +6.483)	(9.401, 10.044)	(31.212, 32.470)	(+3.031, +8.326)	(9.108, 10.086)
HS	40.330	4.211*	16.922	40.560	5.646**	14.385
	(39.205, 41.499)	(−0.146, +8.558)	(16.133, 17.699)	(39.264, 41.908)	(+0.240, +11.250)	(13.405, 15.341)
AA	51.671	1.381	15.379	48.249	−3.261	16.093
	(50.043, 53.251)	(−4.402, +7.132)	(14.443, 16.237)	(46.193, 50.331)	(−10.724, +4.476)	(14.899, 17.181)
BA+	59.370	−0.906	14.915	55.723	4.250	15.253
	(57.678, 61.037)	(−6.153, +4.486)	(14.016, 15.746)	(54.135, 57.345)	(−0.983, +9.441)	(14.391, 16.067)
Men, Urban *Hukou*						
<HS	36.119	4.927**	11.704	37.732	5.045	12.984
	(35.172, 37.123)	(+1.380, +8.575)	(10.856, 12.533)	(35.672, 39.867)	(−3.224, +13.112)	(11.180, 14.592)
HS	42.062	5.406*	14.699	43.875	10.240***	13.346
	(40.902, 43.218)	(+0.886, +9.876)	(13.922, 15.426)	(42.326, 45.447)	(+4.253, +16.453)	(12.229, 14.351)
AA	49.601	−0.911	13.051	48.337	5.114	13.454
	(48.172, 51.92)	(−7.926, +5.788)	(12.130, 13.920)	(46.538, 50.147)	(−3.506, +13.340)	(12.340, 14.479)
BA+	54.503	3.681	14.842	54.418	6.656**	14.030
	(52.902, 56.056)	(−3.304, +11.170)	(14.014, 15.605)	(53.041, 55.798)	(+0.377, +13.175)	(13.266, 14.744)

Significance Level: * p-value < 0.05, ** p-value < 0.01, *** p-value < 0.001

Source: CGSS 2008~2021 Data.

Note: Bootstrapped 95% intervals in parentheses. All γ estimates are statistically significant with p-value < 0.001.

Table A8. Average Parental and Adult Occupation Ranks and their Cross-Cohort Changes by Educational Attainment

	Older Cohort (1964~1979)	Younger Cohort (1980~1995)	Δ $\mu_g^{young} - \mu_g^{old}$
Parental Rank (μ_g^p)			
<HS	.440	.392	–.048
	(.437, .443)	(.384, .400)	(–.056, –.039)
HS	.584	.488	–.096
	(.574, .594)	(.475, .502)	(–.113, –.078)
AA	.657	.562	–.094
	(.643, .671)	(.546, .579)	(–.116, –.073)
BA+	.689	.645	–.044
	(.676, .702)	(.633, .657)	(–.061, –.026)
Adult Rank (μ_g^a)			
<HS	.392	.310	–.082
	(.389, .395)	(.303, .318)	(–.090, –.073)
HS	.640	.500	–.141
	(.631, .650)	(.486, .513)	(–.157, –.124)
AA	.788	.624	–.164
	(.777, .798)	(.609, .639)	(–.182, –.145)
BA+	.848	.732	–.116
	(.841, .855)	(.722, .741)	(–.128, –.105)

Source: CGSS 2008~2021 Data.

Note: Bootstrapped 95% intervals in parentheses. Population means of occupation ranks are 0.5.

Table A9. Model-Simulated Contributions of Educational Processes by the Status of Educational Distribution, with Statuses of Control Variables

		α_g		
Control #	Value	Educ. Upgrading	Ctrl. Variable 1	Ctrl. Variable 2
Origin-Education Association, process #1				
			γ_g	v_g
ctrl 3	–.016	old	young	old
	(–.050, –.018)			
ctrl 4	–.017	old	young	young
	(–.052, +.016)			
ctrl 1	–.020	old	old	old
	(–.054, +.013)			
ctrl 2	–.021	old	old	young
	(–.055, +.013)			
ctrl 3	–.023	young	young	old
	(–.058, +.011)			
ctrl 4	–.024	young	young	young
	(–.057, +.010)			
ctrl 1	–.027	young	old	old
	(–.059, +.005)			
ctrl 2	–.028	young	old	young
	(–.060, +.005)			
Education-Destination Association, process #2				
			λ_g	v_g
ctrl 4	–.017	old	young	young
	(–.052, +.016)			
ctrl 3	–.027	old	young	old
	(–.061, +.007)			
ctrl 2	–.030	old	old	young
	(–.064, +.003)			
ctrl 1	–.031	old	old	old
	(–.065, +.002)			
ctrl 4	–.032	young	young	young
	(–.065, +.001)			
ctrl 3	–.035	young	young	old
	(–.068, –.001)			
ctrl 2	–.036	young	old	young
	(–.069, –.004)			
ctrl 1	–.039	young	old	old

		α_g		
Control #	Value	Educ. Upgrading	Ctrl. Variable 1	Ctrl. Variable 2
	(–.072, –.005)			
Within-Education-Group Gap, process #3				
			λ_g	γ_g
ctrl 2	+.020	young	old	young
	(–.014, +.054)			
ctrl 4	+.019	young	young	young
	(–.014, +.053)			
ctrl 1	+.018	young	old	old
	(–.015, +.050)			
ctrl 3	+.017	young	young	old
	(–.016, +.050)			
ctrl 2	–.003	old	old	young
	(–.037, +.030)			
ctrl 4	–.005	old	young	young
	(–.039, +.029)			
ctrl 1	–.005	old	old	old
	(–.038, +.029)			
ctrl 3	–.006	old	young	old
	(–.039, +.028)			

Source: CGSS 2008~2021 Data.

Note: 95% intervals from repeated simulations in parentheses. The order of rows is consistent with the order of bars in each panel. λ_g represents to control origin-education association (#1); γ_g represents to control education-destination association (#2); and υ_g represents to control within-education-group gaps (#3). The status of variables being "old" represents that they are set to be unchanged, using older cohort's parameters; the status being "young" represents that the variables changed as observed, using younger cohort's parameters. In scenario *control 1*, both control variables are unchanged; in *control 2*, control variable 1 is unchanged while control variable 2 is changed; in *control 3*, control variables 1 and 2 are changed and unchanged, respectively; in *control 4*, both are changed.

Table A10. Contribution of Educational Processes to Cross-Cohort Change in Intergenerational Occupation Persistence by Model Simulations, among Women and Men

	As Only Unchanged	As Only Changed	Full Average of 8 Pairs
Women			
Origin-Education Association	–.041	–.035	–.038
	(–.090, +.007)	(–.083, +.013)	(–.055, –.022)
Education-Destination Association	–.045	–.043	–.044
	(–.092, +.003)	(–.091, +.005)	(–.061, –.027)
Within-Education-Group Gaps	–.001	–.022	–.012
	(–.050, +.047)	(–.069, +.026)	(–.029, +.005)
Shares of Education Group	+.043	+.040	+.041
	(–.008, +.092)	(–.007, +.086)	(+.024, +.058)
Men			
Origin-Education Association	–.010	–.009	–.010
	(–.056, +.036)	(–.055, +.038)	(–.026, +.007)
Education-Destination Association	–.023	–.022	–.023
	(–.069, +.023)	(–.070, +.025)	(–.040, –.007)
Within-Education-Group Gaps	+.038	+.016	+.026
	(–.009, +.085)	(–.031, +.062)	(+.010, +.043)
Shares of Education Group	+.016	+.006	+.011
	(–.031, +.062)	(–.040, +.053)	(–.006, +.027)

Source: CGSS 2008~2021 Data.

Note: 95% intervals from repeated simulations in parentheses. Column 1 treats the target factor as the only unchanged (using old parameters), while other three changed as observed. Column 2 treats the target factor as the only changed (using young parameters). Column 3 takes the average of all 8 possible pairs of comparison. See Appendix III – The Model-Based Approach – Counterfactual Analyses for details.

Table A11. Gender Differences: Overall Intergenerational Occupation Persistence by Cohort and its Cross-Cohort Change

	Older Cohort (1964~1979)	Younger Cohort (1980~1995)	Δ $\beta^{young} - \beta^{old}$
Both *Hukou* Status	+.069	–.011	–.079
	(+.038, +.100)	(–.058, +.038)	(–.137, –.022)
Rural *Hukou*	+.020	+.020	0.000
	(–.021, +.060)	(–.046, +.085)	(–.077, +.078)
Urban *Hukou*	+.030	–.062	–.092
	(–.034, +.096)	(–.172, +.055)	(–.222, +.040)

Source: CGSS 2008~2021 Data.

Note: Bootstrapped 95% intervals in parentheses. Positive numbers represent higher values for women than men; negative numbers represent lower values for women than men. "Both *Hukou* Status" compares all women and all men; "Rural *Hukou*" compares rural women and rural men; and "Urban *Hukou*" compares urban women and urban men.

Table A12. Gender Differences: Contribution of Educational Processes to Cross-Cohort Change in Intergenerational Occupation Persistence by Model Simulations

	As Only Unchanged	As Only Changed	Full Average of 8 Pairs
Both *Hukou* Status			
Origin-Education Association	–.031	–.027	–.029
	(–.099, +.036)	(–.094, +.040)	(–.052, –.005)
Education-Destination Association	–.022	–.021	–.021
	(–.089, +.045)	(–.088, +.047)	(–.045, +.003)
Within-Education-Group Gaps	–.039	–.038	–.038
	(–.106, +.029)	(–.104, +.028)	(–.062, –.015)
Shares of Education Group	+.027	+.033	+.030
	(–.043, +.095)	(–.032, +.099)	(+.007, +.054)

	As Only Unchanged	As Only Changed	Full Average of 8 Pairs
Rural *Hukou*			
Origin-Education Association	+.003	+.009	+.007
	(–.091, +.097)	(–.087, +.106)	(–.026, +.041)
Education-Destination Association	–.011	–.006	–.010
	(–.110, +.077)	(–.100, +.091)	(–.044, +.023)
Within-Education-Group Gaps	–.030	–.036	–.032
	(–.124, +.064)	(–.130, +.059)	(–.065, +.002)
Shares of Education Group	+.038	+.033	+.037
	(–.056, +.132)	(–.061, +.126)	(+.004, +.070)
Urban *Hukou*			
Origin-Education Association	–.028	–.041	–.033
	(–.190, +.135)	(–.202, +.118)	(–.091, +.025)
Education-Destination Association	–.025	–.012	–.018
	(–.186, +.137)	(–.175, +.148)	(–.075, +.040)
Within-Education-Group Gaps	–.062	–.044	–.051
	(–.225, +.099)	(–.203, +.114)	(–.108, +.006)
Shares of Education Group	–.015	+.015	+.001
	(–.178, +.148)	(–.147, +.178)	(–.057, +.059)

Source: CGSS 2008~2021 Data.

Note: 95% intervals from repeated simulations in parentheses. Column 1 treats the target factor as the only *unchanged* (using old parameters), while other three changed as observed. Column 2 treats the target factor as the only *changed* (using young parameters). Column 3 takes the average of all 8 possible pairs of comparison. See Appendix III – The Model-Based Approach – Counterfactual Analyses for details. Positive numbers represent higher values for women than men; negative numbers represent lower values for women than men. "Both *Hukou* Status" compares all women and all men; "Rural *Hukou*" compares rural women and rural men; and "Urban *Hukou*" compares urban women and urban men.

Table A13. Contribution of Educational Processes to Cross-Cohort Change in Intergenerational Occupation Persistence by Model Simulations, in *Hukou*-Specific and Gender-and-*Hukou*-Specific Sub-Samples

	As Only Unchanged	As Only Changed	Full Average of 8 Pairs
Rural *Hukou* Holders			
Origin-Education Association	–.022	–.014	–.018
	(–.069, +.025)	(–.061, +.034)	(–.035, –.002)
Education-Destination Association	–.018	–.015	–.017
	(–.065, +.029)	(–.063, +.032)	(–.034, 0.000)
Within-Education-Group Gaps	+.005	–.001	+.002
	(–.041, +.052)	(–.049, +.046)	(–.015, +.019)
Shares of Education Group	+.023	+.035	+.028
	(–.025, +.071)	(–.012, +.082)	(+.012, +.045)
Urban *Hukou* Holders			
Origin-Education Association	+.020	+.027	+.024
	(–.058, +.098)	(–.051, +.104)	(–.004, +.051)
Education-Destination Association	–.021	–.016	–.018
	(–.099, +.057)	(–.095, +.062)	(–.046, +.010)
Within-Education-Group Gaps	+.036	–.031	+.034
	(–.044, +.114)	(–.047, +.108)	(+.006, +.061)
Shares of Education Group	–.015	–.014	–.014
	(–.094, +.063)	(–.093, +.065)	(+.041, +.014)
Women, Rural *Hukou*			
Origin-Education Association	–.021	–.010	–.015
	(–.086, +.045)	(–.077, +.057)	(–.038, +.009)
Education-Destination Association	–.028	–.018	–.023
	(–.092, +.037)	(–.085, +.050)	(–.046, 0.000)
Within-Education-Group Gaps	–.011	–.018	–.014
	(–.077, +.055)	(–.085, +.050)	(–.037, +.010)
Shares of Education Group	+.045	+.057	+.051
	(–.020, +.113)	(–.009, +.123)	(+.028, +.075)
Women, Urban *Hukou*			
Origin-Education Association	+.002	+.007	+.005
	(–.135, +.139)	(–.126, +.136)	(–.042, +.053)
Education-Destination Association	–.038	–.024	–.030
	(–.170, +.095)	(–.155, +.107)	(–.077, +.017)
Within-Education-Group Gaps	+.004	–.010	+.008
	(–.131, +.138)	(–.122, +.139)	(–.039, +.055)

	As Only Unchanged	As Only Changed	Full Average of 8 Pairs
Shares of Education Group	−.021	−.007	−.013
	(−.155, +.114)	(−.137, +.126)	(−.060, +.035)
Men, Rural *Hukou*			
Origin-Education Association	−.024	−.020	−.023
	(−.092, +.044)	(−.088, +.047)	(−.046, +.001)
Education-Destination Association	−.012	−.012	−.013
	(−.079, +.055)	(−.080, +.054)	(−.036, +.011)
Within-Education-Group Gaps	+.019	+.018	+.018
	(−.048, +.085)	(−.049, +.083)	(−.006, +.042)
Shares of Education Group	+.007	+.024	+.014
	(−.061, +.074)	(−.043, +.090)	(−.009, +.038)
Men, Urban *Hukou*			
Origin-Education Association	+.030	+.048	+.039
	(−.062, +.123)	(−.046, +.141)	(+.006, +.072)
Education-Destination Association	−.013	−.012	−.012
	(−.104, +.078)	(−.106, +.085)	(−.045, +.021)
Within-Education-Group Gaps	+.066	+.053	+.059
	(−.027, +.159)	(−.040, +.148)	(+.026, +.093)
Shares of Education Group	−.007	−.021	−.014
	(−.099, +.085)	(−.118, +.074)	(−.048, +.019)

Source: CGSS 2008~2021 Data.

Note: 95% intervals from repeated simulations in parentheses. Column 1 treats the target factor as the only unchanged (using old parameters), while other three changed as observed. Column 2 treats the target factor as the only changed (using young parameters). Column 3 takes the average of all 8 possible pairs of comparison. See Appendix III – The Model-Based Approach – Counterfactual Analyses for details.

Table A14. Rural-Urban (*Hukou*) Differences: Overall Intergenerational Occupation Persistence by Cohort and its Cross-Cohort Change

	Older Cohort (1964~1979)	Younger Cohort (1980~1995)	Δ $\beta^{young} - \beta^{old}$
Both Gender	+.024	–.026	–.049
	(–.014, +.061)	(–.089, +.036)	(–.123, +.022)
Men	+.020	+.020	–.001
	(–.037, +.078)	(–.087, +.117)	(–.122, +.114)
Women	+.030	–.062	–.092
	(–.021, +.080)	(–.142, +.018)	(–.188 +.002)

Source: CGSS 2008~2021 Data.

Note: Bootstrapped 95% intervals in parentheses. Positive numbers represent higher values for rural *hukou* than urban *hukou*; negative numbers represent lower values for rural *hukou* than urban *hukou*. "Both Gender" compares all rural *hukou* holders and all urban *hukou* holders; "Women" compares women with rural *hukou* and women with urban *hukou*; and "Men" compares men with rural *hukou* and men with urban *hukou*.

Table A15. Rural-Urban (*Hukou*) Differences: Contribution of Educational Processes to Cross-Cohort Change in Intergenerational Occupation Persistence by Model Simulations

	As Only Unchanged	As Only Changed	Full Average of 8 Pairs
Both Gender			
Origin-Education Association	–.042	–.041	–.042
	(–.133, +.049)	(–.131, +.051)	(–.074, –.010)
Education-Destination Association	+.002	+.001	+.001
	(–.088, +.094)	(–.090, +.094)	(–.033, +.033)
Within-Education-Group Gaps	–.030	–.032	–.032
	(–.122, +.061)	(–.122, +.059)	(–.064, 0.000)
Shares of Education Group	+.038	+.049	+.042
	(–.053, +.130)	(–.042, +.141)	(+.010, +.075)

	As Only Unchanged	As Only Changed	Full Average of 8 Pairs
Women			
Origin-Education Association	–.023	–.017	–.020
	(–.174, +.129)	(–.161, +.130)	(–.074, +.032)
Education-Destination Association	+.010	+.006	+.007
	(–.138, +.157)	(–.142, +.154)	(–.045, +.059)
Within-Education-Group Gaps	–.015	–.028	–.022
	(–.164, +.136)	(–.173, +.120)	(–.074, +.032)
Shares of Education Group	+.066	+.064	+.064
	(–.081, +.214)	(–.083, +.211)	(+.011, +.117)
Men			
Origin-Education Association	–.054	–.067	–.061
	(–.170, +.060)	(–.182, +.048)	(–.102, –.021)
Education-Destination Association	+.001	–.001	–.001
	(–.113, +.115)	(–.118, +.116)	(–.041, +.040)
Within-Education-Group Gaps	–.047	–.036	–.041
	(–.162, +.068)	(–.150, +.080)	(–.082, 0.000)
Shares of Education Group	+.014	+.045	+.029
	(–.100, +.129)	(–.072, +.162)	(–.012, +.069)

Source: CGSS 2008~2021 Data.

Note: 95% intervals from repeated simulations in parentheses. Column 1 treats the target factor as the only *unchanged* (using old parameters), while other three changed as observed. Column 2 treats the target factor as the only *changed* (using young parameters). Column 3 takes the average of all 8 possible pairs of comparison. See Appendix III – The Model-Based Approach – Counterfactual Analyses for details. Positive numbers represent higher values for rural *hukou* than urban *hukou*; negative numbers represent lower values for rural *hukou* than urban *hukou*. "Both Gender" compares all rural *hukou* holders and all urban *hukou* holders; "Women" compares women with rural *hukou* and women with urban *hukou*; and "Men" compares men with rural *hukou* and men with urban *hukou*.

APPENDIX II
The Approach of Mathematical Decomposition

Analytic Approach

An alternative approach to investigate the contributions of educational processes is the mathematical decomposition, which isolates changes in persistence due to changes in between-education-group inequalities (the joint contribution of [1] origin-education and [2] education-destination associations), (3) within-education-group gaps, and (4) changes in educational composition. Changes in the first two educational processes – origin-education and education-destination associations – are estimated as a whole and isolated as one component by this approach. The analyses will first partition persistence within each cohort into pieces reflecting each of these components, and then examine the changes in and the contribution of these components across cohorts.

$$\beta = \sum_{g} \pi_g \cdot B_g \tag{A1}$$

Within each gender and cohort, the rank-rank slope β is the weighted average of the "total associations" (more details described later) between the parental and adult occupational ranks within education groups (Equation A1). The share of education group g (for instance, the percentage of people holding a bachelor's degree or above), π_g, serves as the weight. The more prevalent a given education group, the larger its impact on the overall rank-rank slope β. If the share of an education group (for example, college graduates) increased after the educational expansion, then its π_g value should be larger

after the expansion than before, and its impact on β also consequently grows. Later, such changes in π_g values will be used to isolate the role in persistence trends of changes in the educational composition. The total association within education group g is recorded as B_g, and it is the sum of the between-group association and the within-group association. It is further decomposed in the next steps.

$$B_g = b_g + w_g \tag{A2}$$

In Equation A2, I split the total association between parents' and adults' occupational ranks in education group g, B_g, into the between-group association b_g and the within-group association w_g (also see Torche and Corvalan 2018).[i] This step enables me to isolate the role of origin-education associations and education-destination associations together, as well as the role of the differences in occupational attainment by social origin within education groups.

$$b_g = \frac{\left(\mu_g^p - \mu^p\right) \cdot \left(\mu_g^a - \mu^a\right)}{\left(\sigma^p\right)^2} \tag{A3}$$

The between-group association (Equation A3), b_g, captures the gaps in occupation statuses between education groups in childhood (reflecting the origin-education associations) and adulthood (reflecting education-destination associations). The wider

[i] The between-group association (b_g) is necessary because, when I run regressions within education groups, the slopes (B_g values) reflect regression to group-specific averages. Since these averages differ, these slopes cannot capture all inequalities that contribute to the intergenerational occupation correlation in the full population.

these gaps between the average parental and adult occupation statuses in education groups, and the average statuses of the entire population, the larger the b_g values. For instance, high school dropouts tend to come from lower social origins (where the parents' jobs are of lower status), and arrive in lower social destinations (where the adults' jobs are of lower status), than their more educated peers. Consequently, their between-group association is quite high, because their typical occupation status is far from (lower than) the overall average. On the other hand, since the origin and destination statuses of college graduates tend to be higher, and thus being higher than the overall average, their between-group association is higher as well. Because the between-group association makes up a part of the total association, all else being equal, a larger between-group association indicates a larger total association.

Formally, this between-group association b_g consists of the contribution of education group g to the between-group covariance between parents' and children's occupational ranks, $\left(\mu_g^p - \mu^p\right) \cdot \left(\mu_g^a - \mu^a\right)$, scaled by the variance of parental occupational ranks, $(\sigma^p)^2$. These are, respectively, the nominator and the denominator on the right of Equation A3.

Substantively, the term $\mu_g^p - \mu^p$ captures the origin-education association, where μ_g^p is the average parental occupational rank in education group g and μ^p is the overall average parental occupational rank. When the average parental rank in an

education group (e.g., among high school dropouts or college graduates) is farther from the overall average, the difference $\mu_g^p - \mu^p$ is larger, as is the between-group association. A larger difference $\mu_g^p - \mu^p$ indicates a greater origin-education association.

The other term, $\mu_g^a - \mu^a$, captures the education-destination association, where μ_g^a is the average adult occupational rank in education group g and μ^a is the overall average adult occupational rank. Similarly, the farther the average adult rank in a group from the overall average, the larger the difference $\mu_g^a - \mu^a$, and the larger the between-group association. The education-destination association is greater when this difference $\mu_g^a - \mu^a$ is larger. In summary, when there are larger differences between the averages of parent and adult occupational ranks in an education group and the overall averages, the between-group association is larger. This indicates a higher intergenerational occupation persistence.

$$w_g = \beta_g^w \cdot \frac{\left(\sigma_g^p\right)^2}{\left(\sigma^p\right)^2} \tag{A4}$$

The within-group association, w_g, captures how strongly occupational rank persists across generations among people in the education group g. When the gaps in occupational outcomes by social origin are smaller within an education group (e.g., college graduates) compared to other groups, its within-group association should also be smaller. If the gaps are widened, w_g should also increase. In addition, if the w_g values of college graduates are smaller than those of non-graduates in both cohorts, the

increased percentages of college graduates due to the educational expansion should be able to reduce the intergenerational occupation persistence. A larger within-group association indicates that there is a larger gap in occupation outcomes by origin status among people within education group g. Subsequently, as another part of the total association, all else being equal, a larger within-group association indicates a larger total association.

w_g is calculated according to Equation A4, having the within-education-group rank-rank slope, β_g^w, scaled by the quotient of $\left(\sigma_g^p\right)^2$, the variance of parental occupational rank in education group g, by $(\sigma^p)^2$, the overall variance of parental occupational rank. The scaling factor, $\frac{\left(\sigma_g^p\right)^2}{(\sigma^p)^2}$, ensures that the weighted sum of the between- and within-group associations equals the overall rank-rank slope.[ii]

$$O_{gi}^a = \alpha_g + \beta_g^w \cdot O_{gi}^p + \epsilon_{gi} \tag{A5}$$

Specifically, the within-education-group rank-rank slope, β_g^w, is calculated according to Equation A5. It is exactly the same as Equation 1, except that individuals are categorized into education groups rather than pooled together. Within education group g, O_{gi}^a is the adult occupational rank of observation i; their parental occupational rank is recorded as O_{gi}^p; and ϵ_{gi} is the error. In education group g, the smaller the

[ii] All variances in this decomposition are calculated using a population variance estimator, dividing the sum of the squared deviations from the mean by the sample size N, rather than the typically used divisor to adjust for the loss in degrees of freedom from estimating the mean $N-1$ (Bloome et al. 2018:1247, Note#16).

occupational attainment gaps by origin status, the smaller the within-education-group rank-rank slope β_g^w. For example, since I hypothesize that the within-education-group gap among less-educated people remained stable, their β_g^w value should stay generally unchanged across cohorts; for highly educated individuals, I hypothesize that the gaps grew, increasing β_g^w across cohorts. Moreover, since the within-education-group gaps in social destination by social origin among college graduates should be smaller than those among their counterparts with less education, college graduates' β_g^w should be smaller for groups with more education than those with less.

$$\beta = \sum_g \pi_g \cdot B_g \tag{A6.1}$$

$$= \sum_g \pi_g \cdot \left(b_g + w_g\right) \tag{A6.2}$$

$$= \sum_g \pi_g \cdot \left[\frac{\left(\mu_g^p - \mu^p\right) \cdot \left(\mu_g^a - \mu^a\right)}{\left(\sigma^p\right)^2} + \beta_g^w \cdot \frac{\left(\sigma_g^p\right)^2}{\left(\sigma^p\right)^2}\right] \tag{A6.3}$$

Putting Equations A1-A5 together, Equation A6 shows how the overall rank-rank slope β is decomposed and how its components are calculated. The point estimates for all population quantities, such as $\hat{\pi}_g$, $\hat{\beta}_g$, $\bar{x}_g^a$, and $\hat{\sigma}_g^p$, are generated from the sample data with the two genders and two cohorts separated.

Together, changes in these components capture the roles of the four educational processes in intergenerational occupation persistence within each cohort, including the between-group inequalities (the joint contribution of [1] origin-education associations

and [2] education-destination associations together, b_g, expected to increase persistence), (3) within-group gaps (in occupational attainment by social origin within education groups, w_g, expected to increase persistence), and (4) educational composition (π_g, expected to reduce persistence). (See Appendix III – Mathematical Decomposition – The Equation for details.)

$$\beta^{young} - \beta^{old} = \underbrace{\sum_g \pi_g^{old} \cdot \left(b_g^{young} - b_g^{old}\right)}_{\text{Between-Group Inequalities}} + \underbrace{\sum_g \pi_g^{old} \cdot \left(w_g^{young} - w_g^{old}\right)}_{\text{Within-Group Gaps}} + \underbrace{\sum_g B_g^{young} \cdot \left(\pi_g^{young} - \pi_g^{old}\right)}_{\text{Educational Composition}} \quad \text{(A7)}$$

The contribution of changes in these components to the cross-cohort change in intergenerational occupation persistence, $\beta^{young} - \beta^{old}$, are calculated according to Equation 8. β^{old} and β^{young} are the overall rank-rank slopes for people in the older cohort (1964-1979) and those in the younger cohort (1980-1995), respectively. Each one of these three summations captures the contribution of one component, respectively. Each component is a weighted average of cross-cohort changes, weighted by, for the first two components, the education groups' proportions in the older cohort (π_g^{old}) and, for the last component, by the education groups' total associations in the younger cohort (B_g^{young}). Thus, the weights, the magnitudes of the differences, and their signs together

determine the role of the components. Since the quantities of interest are non-linear combinations of the component statistics, I assess sampling uncertainty via the nonparametric bootstrap.[iii]

Results

The approach of decomposition relies on, for each education group, three components of intergenerational persistence – "between-education-group" association that captures the between-education group inequalities, "within-education-group" association that captures gaps in destination status by origin status within education groups, and "% Share" of education groups that captures the educational distribution. These three items are recorded in Table A16. Specifically, the second column is the scaled version of group-specific rank-rank slopes (Table A4), and the third column is equivalent to Table A5.

For both the between-education-group and the within-education-group associations, different groups experienced different changes. Therefore, their contributions to persistence trends will depend on the specific magnitudes of association changes and the prevalence of education groups (first and second terms in Equation A7). Since the group

[iii] I resample with replacement to obtain 21,120 bootstrap samples; for each bootstrap sample, I calculate the occupation ranks and quantities of interest. I take the 2.5th and the 97.5th percentiles of the resampling distribution to estimate the 95% intervals.

of high school dropouts was by far the biggest in the older cohort (~69%), therefore carrying more weight in the overall effect, its pattern tends to generally dominate the role of between-education-group inequalities. The contribution of educational upgrading (increasing percentages of college graduates) is expected to be decreasing persistence (Breen 2010:364).

Table A16. Components of Intergenerational Occupation Persistence by Educational Attainment and Cohort (Age Group) and their Cross-Cohort Changes

	Point Estimates			95% Intervals		
	Between Association	Within Association	% Share	Between Association	Within Association	% Share
Older Cohort (1964~1979 group)						
<HS	.091	.136	68.8	(.085, .098)	(.121, .152)	(68.1, 69.5)
HS	.166	.125	14.1	(.142, .191)	(.086, .165)	(13.6, 14.6)
AA	.635	.014	8.2	(.575, .696)	(−.029, .055)	(7.7, 8.6)
BA+	.925	.017	8.9	(.855, .994)	(−.011, .046)	(8.4, 9.3)
Younger Cohort (1980~1995 group)						
<HS	.263	.098	40.7	(.239, .287)	(.070, .125)	(39.4, 42.0)
HS	.000	.097	18.7	(−.002, .003)	(.046, .147)	(17.7, 19.7)
AA	.099	.028	15.7	(.071, .130)	(−.027, .083)	(14.8, 16.6)
BA+	.432	.071	24.9	(.391, .476)	(.032, .111)	(23.8, 26.0)
Cross-Cohort Change						
<HS	+.171	−.039	−28.2	(+.147, +.197)	(−.070, −.007)	(−29.6, −26.7)
HS	−.166	−.029	+4.6	(−.191, −.142)	(−.093, +.035)	(+3.5, +5.7)
AA	−.536	+.014	+7.5	(−.603, −.496)	(−.056, +.084)	(+6.5, +8.5)
BA+	−.493	+.054	+16.1	(−.572, −.410)	(+.005, +.103)	(+14.9, +17.2)

Source: CGSS 2008~2021 Data.

Note: Bootstrapped 95% intervals in parentheses. The "between association," b_g, captures the in-equalities between education groups; the "within association," w_g, captures the gaps in social destination by social origin within education groups; and the "% share," denoted by π_g, is the percentage of the education groups in the cohort. Positive cross-cohort changes represent that the values increased for the younger cohort.

I formally decomposed the cross-cohort changes in the intergenerational occupation persistence of women and men in Table A17, quantifying: the role of changes in (1) origin-education association and (2) education-destination association together (i.e., between-education-group inequalities), (3) gaps in destination status by origin status within education groups, and (4) the educational distribution. Simply put, stability in overall intergenerational persistence is due to offsetting influences of the four educational processes.

Table A17. Decomposition of Cross-Cohort Change in Intergenerational Occupation Persistence

	Δ Point Estimate	Δ 95% Interval
Total	–.025	(–.053, +.004)
Due to Δ between-group inequalities	+.007	(–.014, +.028)
Due to Δ within-group destination gaps	–.025	(–.050, 0.000)
Due to Δ educational composition	–.007	(–.021, +.007)

Source: CGSS 2008~2021 Data.

Note: Bootstrapped 95% intervals in parentheses (second column).

Between-Group Inequalities

As shown in Table A17, the changing between-education-group inequalities overall increased intergenerational persistence by 0.007, all else being equal. This value is the result of the offsetting contributions from different parts of the education distribution described above. While the "between association" values for education groups other than high school dropouts declined, the group for which it increased – high school dropouts – was by far the biggest group (~69%) and therefore carries more weight in the overall

effect. Therefore, though it was the only source of upward force, the group contribution of high school dropouts was strong enough to cancel out the downward pressure from groups with higher education and push the overall contribution toward a positive value. As a result, changes in the origin-education association and the education-destination association together increased intergenerational occupation persistence.

Within-Group Gaps

Similarly, the overall contribution of changing gaps in social destination by social origin within education groups is also determined by offsetting contributions from different groups Together, changes in these within-education-group gaps decreased intergenerational persistence by .025. This negative value resulted from two groups exerting downward pressure on persistence – high school dropouts and graduates – for whom the gaps narrowed in the post-expansion period and who jointly comprised 83 percent of the older cohorts. Consequently, the total contribution of within-education-group gaps in social destination by social origin was pushed toward a negative value, contributing to the decline of intergenerational occupation persistence.

Educational Composition

Lastly, changes in the educational composition, i.e., shares of education groups, have also decreased intergenerational occupation persistence (–.007). This is consistent with our expectation given the changing educational distribution shown above, which shows

that the post-expansion period shifted individuals out of the high-persistence group of high school dropouts to the lower-persistence groups of high school and college graduates.

From Educational Processes to Overall Persistence

Putting these estimates together, Table A17 shows that the slight and statistically insignificant decline in overall intergenerational persistence resulted from offsetting influences of the four educational processes. If the only change across cohorts had been the strengthened origin-education and education-destination association, persistence would have increased by .007. However, the growing percentages of people with higher education (changes in educational composition, last row of Table A17) offset this increase. In addition, changing gaps in people's social destination by social origin within education groups exerted another downward force, leading to the overall decline (though statistically insignificant) in persistence.

Based on the 95% confidence intervals, however, there is insufficient statistical evidence to confirm the roles of the educational processes. All three confidence intervals overlap zero, indicating that we don't have enough evidence to reject the null hypothesis that there are no contributions of educational processes. This does not contradict the point estimates above, but suggests that further explorations are necessary to truly

understand the multifaceted roles of education in intergenerational occupation persistence in postsocialist China.

Robustness Check

I also ran the same analyses with gender-specific, *hukou*-specific, and gender-and-*hukou*-specific samples as robustness check. The estimated contributions are in Tables A18, A19, and A20, respectively. The point estimates also seem to support that the overall stability in intergenerational persistence resulted from offsetting contributions of the educational trends. However, again, there was not enough statistical evidence for me to confirm these contributions. There were not statistically significant differences by gender and *hukou*, either (see Tables A21 and A22). (Tables A18-A22 are at the end of Appendix II.)

The Approaches of Mathematical Decomposition and Model Simulation

Theoretically, while results from the approach of decomposition can sum up and return the overall persistence, it has several limitations that make me decided not to use it in the text as well but leave it in the appendix. Firstly, the roles of origin-education and education-destination associations (process #1 and #2) are estimated together by decomposition, while the model simulations detangle the two. In addition, the approach of model simulation relaxes the assumption that changes in the shares of education

groups (π_g) would leave the other quantities in the population, such as the group averages of parental and adult occupation ranks (μ_g^p and μ_g^a), unchanged, allowing changes in one educational process to alter the roles of another. Lastly, the model simulations further relax the assumptions of linearity and additivity, the details of which are introduced in the text. Therefore, even though the results from model simulations cannot be added up together to determine the overall persistence, I applied the approach of model simulations in the text given these strengths.

The actual results from the decomposition seem to support our expectations based on the descriptive patterns in general, which are different from the model-simulated estimates. I proposed a potential explanation of this discrepancy in the sections of results (with Figure 5 and Table 3) and discussion.

Table A18. Decomposition of Cross-Cohort Change in Intergenerational Occupation Persistence, among Women and Men

	Δ Women	Δ Men
Total	−.067	+.013
	(−.106, −.027)	(−.029, +.053)
Due to Δ between-education-group inequalities	−.016	+.024
	(−.047, +.016)	(−.004, +.054)
Due to Δ within-education-group gaps	−.035	−.010
	(−.070, +.001)	(−.044, +.024)
Due to Δ educational composition	−.016	−.001
	(−.038, +.005)	(−.020, +.016)

Source: CGSS 2008~2021 Data.

Note: Bootstrapped 95% intervals in parentheses.

Table A19. Decomposition of Cross-Cohort Change in Intergenerational Occupation Persistence, by *Hukou*

	Δ Rural *Hukou*	Δ Urban *Hukou*
Total	+.007	+.056
	(–.032, +.046)	(–.004, +.119)
Due to Δ between-education-group inequalities	–.002	+.106
	(–.024, +.020)	(+.054, +.163)
Due to Δ within-education-group gaps	–.013	+.040
	(–.045, +.020)	(–.023, +.098)
Due to Δ educational composition	+.022	–.089
	(+.004, +.039)	(–.131, –.045)

Source: CGSS 2008~2021 Data.

Note: Bootstrapped 95% intervals in parentheses. Values are estimated with both genders pooled together.

Table A20. Decomposition of Cross-Cohort Change in Intergenerational Occupation Persistence, by Gender and *Hukou*

	Δ Women	Δ Men
Rural *Hukou*		
Total	+.008	+.008
	(–.045, +.062)	(–.049, +.064)
Due to Δ between-education-group inequalities	+.008	–.011
	(–.021, +.038)	(–.042, +.020)
Due to Δ within-education-group gaps	–.024	0.000
	(–.069, +.024)	(–.045, +.045)
Due to Δ educational composition	+.024	+.019
	(–.003, +.049)	(–.004, +.041)

	Δ Women	Δ Men
Urban *Hukou*		
Total	+.008	+.100
	(–.092, +.118)	(+.024, +.178)
Due to Δ between-education-group inequalities	+.086	+.123
	(–.010, +.184)	(+.066, +.186)
Due to Δ within-education-group gaps	+.013	+.063
	(–.089, +.091)	(–.019, +.143)
Due to Δ educational composition	–.090	–.085
	(–.148, –.007)	(–.143, –.034)

Source: CGSS 2008~2021 Data.

Note: Bootstrapped 95% intervals in parentheses.

Table A21. Gender Differences: Decomposition of Cross-Cohort Change in Intergenerational Occupation Persistence

	Both *Hukou* Status	Rural *Hukou*	Urban *Hukou*
Total	–.079	0.000	–.092
	(–.137, –.022)	(–.077, +.078)	(–.222, +.040)
Due to Δ between-education-group inequalities	–.040	+.019	–.037
	(–.083, +.003)	(–.023, +.061)	(–.154, +.077)
Due to Δ within-education-group gaps	–.024	–.024	–.050
	(–.073, +.024)	(–.087, +.040)	(–.181, +.059)
Due to Δ educational composition	–.014	+.005	–.005
	(–.043, +.014)	(–.029, +.039)	(–.081, +.098)

Source: CGSS 2008~2021 Data.

Note: Bootstrapped 95% intervals in parentheses. Positive numbers represent higher values for women than men; negative numbers represent lower values for women than men. "Both *Hukou* Status" compares all women and all men; "Rural *Hukou*" compares rural women and rural men; and "Urban *Hukou*" compares urban women and urban men.

Table A22. Rural-Urban (*Hukou*) Differences: Decomposition of Cross-Cohort Change in Intergenerational Occupation Persistence

	Both Gender	Women	Men
Total	–.049	–.001	–.092
	(–.123, +.022)	(–.122, +.114)	(–.188 +.002)
Due to Δ between-education-group inequalities	–.108	–.078	–.134
	(–.169, –.151)	(–.181, +.022)	(–.204, –.069)
Due to Δ within-education-group gaps	–.052	–.036	–.063
	(–.118, +.018)	(–.126, +.077)	(–.153, +.030)
Due to Δ educational composition	+.111	+.113	+.104
	(+.064, +.156)	(+.025, +.177)	(+.049, +.166)

Source: CGSS 2008~2021 Data.

Note: Bootstrapped 95% intervals in parentheses. Positive numbers represent higher values for rural *hukou* than urban *hukou*; negative numbers represent lower values for rural *hukou* than urban *hukou*. "Both Gender" compares all rural *hukou* holders and all urban *hukou* holders; "Women" compares women with rural *hukou* and women with urban *hukou*; and "Men" compares men with rural *hukou* and men with urban *hukou*.

APPENDIX III
Additional Descriptions of Analytic Approach with Equations

In this appendix, I use equations to explicitly express my operations and/or provide more details, so that readers can understand my analytic approach better.

Overall Persistence without and with Age Controlled

In Footnote #8, I explained in words how I can control age when estimating the overall intergenerational occupation persistence. The operations are captured by Equations A8 and A9.

$$O_i^a = \gamma_0 + \gamma_1 \cdot age_i + \gamma_2 \cdot age_i^2 + u_i \tag{A8}$$

$$u_i = \gamma_3 + \gamma_4 \cdot O_i^p + w_i \tag{A9}$$

where, for observation i, O_i^a is the adult occupational rank, O_i^p is the parental occupational rank, and age is the age at the time of survey. The ages of observations are all integers, determined by subtracting the year of birth from the year of survey. For example, the age of an individual surveyed in 2012 and born in 1980 is $2012 - 1980 = 32$. In addition, when predicting O_i^a (adult rank) with age and age^2 in Equation A8, γ_0 is the intercept, γ_1 and γ_2 are the slopes respectively associated with age and age^2, and u_i is the residual. Next, in Equation A9 where I predict the u_i with O_i^p (parental rank), γ_3 is the intercept, γ_4 is the associated slope of O_i^p, and w_i is the residual.

The variable of interest is γ_4, in Equation A9. Its values are in the "Age Controlled" rows in Appendix I – Table A1. Compared to the values in the "Simple" rows, which are

the estimates of β in Equation 1, γ_4 values are similar. Therefore, for simplicity, I do not control for age when estimating the overall intergenerational persistence or conducting model simulations.

The Model-Based Approach

Persistence Simulation

Given the estimates of the four parameters (λ_g for origin-education association, γ_g for education-destination association, β_g and v_g for within-education-group gaps, and α_g for educational composition), there are two steps to predict the ISEI scores of individuals' occupations in adulthood, which will then be ranked and used to simulate persistence (in the same way as Equation 1). Therefore, these two steps to predict ISEI are the key of simulate persistence with the models.

The first step relies on the generalized ordered logit model (Equation 2). Given one's parental occupational rank, this step predicts one's educational attainment. As shown in Equation 4, the model returns the probabilities, for respondent i, of obtaining each level of education (E_i). In my thesis, specifically, there are four probabilities – $P(E_i = <\text{HS})$, $P(E_i = \text{HS})$, $P(E_i = \text{AA})$, and $P(E_i = \text{BA+})$ as the probability for being high school dropout, being high school graduates, receiving an associate's degree, and holding a bachelor's degree or above, respectively. For each respondent, their educational

attainment is independently simulated by "rolling a four-sided dice" with different probabilities for each of the four faces.

Suppose there is an individual j, for whom

$$P(E_j = \text{<HS}) = 0.40$$

$$P(E_j = \text{HS}) = 0.30$$

$$P(E_j = \text{AA}) = 0.20$$

$$P(E_j = \text{BA+}) = 0.10$$

Therefore, when individual j roll the dice and receive a "<HS," the probability of which is 0.40, I predict that $E_j = \text{<HS}$, that individual j is a high school dropout. Similarly, if the dice gives individual j a "BA+" (probability 0.10), I predict that $E_j = \text{BA+}$. Operationally, I generate a random number, w_j^*, such that

$$w_j^* \sim \text{Continuous Uniform } (0,1)$$

In other words, w_j^* follows a standard continuous uniform distribution. Then, I determine the educational attainment of individual j based on Equation A8 below.

$$E_j = \begin{cases} \text{<HS} & w_j^* \in [0.00,0.40) \\ \text{HS} & w_j^* \in [0.40,0.70) \\ \text{AA} & w_j^* \in [0.70,0.90) \\ \text{BA+} & w_j^* \in [0.90,1.00] \end{cases} \tag{A10}$$

Therefore, the general form of predicting one's education, E_i, given that

$$P(E_j = \text{<HS}) = a$$

$$P(E_j = \text{HS}) = b$$

$$P(E_j = \text{AA}) = c$$

$$P(E_j = \text{BA+}) = d$$

where $a + b + c + d = 1$, is described in Equation A9,

$$E_j = \begin{cases} <\text{HS} & w_j^* \in [0, a) \\ \text{HS} & w_j^* \in [a, a+b) \\ \text{AA} & w_j^* \in [a+b, a+b+c) \\ \text{BA+} & w_j^* \in [a+b+c, 1] \end{cases} \tag{A11}$$

The second step relies on the normal linear model (Equation 3). Given one's parental occupational rank, this step predicts the ISEI score of one's occupation in adulthood, conditional on the predicted educational attainment. If one was predicted to be high school dropouts in the first step, parameters of $<\text{HS}$ are used in this step; similarly, if one was predicted to hold a bachelor's degree or above, parameters of BA+ are used. The calculation is described as Equation 6. Specifically, the multiplier of randomness, z_i^*, may sound vague in words (the inverse of the standard normal cumulative distribution function, evaluated at a randomly generated value under a standard uniform distribution, for observation i). Let me say it in this way:

Recall that, for a continuous random variable X such that

$$X \sim \text{Normal}\left(\mu, \sigma^2\right)$$

where μ is the mean and σ is the standard deviation, the probability density function is

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

Then, if X follows a *standard* normal distribution, i.e.,

$$X \sim \text{Normal } (0,1)$$

we would have the probability density function

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$$

Next, I generate a random number (say, p_i^*) that also follows a standard uniform distribution, such that

$$P[X \leq z_i^*] = \int_{-\infty}^{z_i^*} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}\, dx = p_i^* \tag{A12}$$

Note that $P[X \leq z_i^*]$ can also be determined using the cumulative density function of X, $F_X(x)$.

Therefore, put together, z_i^* is calculated based on Equation A10,

$$F_X(z_i^*) = p_i^* \tag{A13}$$

where X is a continuous random variable following a standard normal distribution and p_i^* follows a standard uniform distribution.

Counterfactual Analyses

The counterfactual analyses simulate persistence in the same way as above, but they use different combinations of parameters from the models (λ_g for origin-education association, γ_g for education-destination association, β_g and υ_g for within-education-group gaps, and α_g for educational composition) to reflect different counterfactual

scenarios. To represent these combinations, I use different subscripts[iv] for persistence β, our key factor of interest. When a specific parameter is present in the subscript, persistence is simulated as if this parameter has changed (using younger cohorts' estimates); when it is missing in the subscript, I estimate persistence as if this parameter has remained unchanged (using older cohorts' estimates).

For example,

$$\beta_{\lambda\gamma\upsilon\alpha}$$

represent that, in this scenario, all four parameters have changed (using younger cohorts' estimates), because all is present in the subscript. This is exactly the same as the simulated persistence of the younger cohort. Also,

$$\beta_{\lambda\gamma\upsilon}$$

represent that the origin-education and education-destination associations, as well as the within-education-group gap, have changed (younger cohorts' estimates), while the shares of education groups (process #4) have remained unchanged (older cohorts' estimates), because α is missing in the subscript. Next, if the origin-education and education-destination associations (process #1 and #2) remained unchanged and the other two have changed, I use

[iv] The parameters for process #3, within-education-group gaps, are β and υ. They change together – for both, I use the estimates from either the older cohort or from the younger cohort; I never use, say, β from the older cohort and υ from the younger cohort. Therefore, since β_β looks very confusing, I use υ in the subscripts to show the status.

$$\beta_{v\alpha}$$

to represent the scenario. Lastly, when everything remained unchanged, the scenario is shown by

$$\beta_{(1)}$$

where the parentheses around the number 1 is to distinguish it from β_1, which usually represent something else in regressions. Note that this value is simulated by applying all four sets of parameters estimated from the older cohort *on the younger cohort*, so it is *not* the older cohorts' simulated persistence.

Given this type of notations, I can represent the three ways to estimate the roles of educational processes (Table 2), as well as the three ways to test the existence of interactions (Table 3), easily. Suppose we want to test the role of the origin-education expansion, captured by λ.

Column 1 in Table 2 treats the target factor (λ) as the only unchanged, while other three changed as observed. It relies on

$$\Delta_1 = \beta_{\lambda\gamma v\alpha} - \beta_{\gamma v\alpha}$$

and tests

$$H_0: \Delta_1 = 0$$

$$H_a: \Delta_1 \neq 0$$

Column 2 in Table 2 treats the target factor (λ) as the only changed, relying on

$$\Delta_2 = \beta_\lambda - \beta_{(1)}$$

and tests

$$H_0: \Delta_2 = 0$$

$$H_a: \Delta_2 \neq 0$$

Column 3 takes all 16 counterfactual scenarios into account, using the average from 8 pairs of comparisons as the result. These 8 pairs come from the combinations of the other three parameters to control, each of which has two statuses (so $2^3 = 8$). The result is

$$\Delta_3 = \frac{1}{8}\left[\left(\beta_{\lambda\gamma\upsilon\alpha} - \beta_{\gamma\upsilon\alpha}\right) + \left(\beta_{\lambda\upsilon\alpha} - \beta_{\upsilon\alpha}\right) + \left(\beta_{\lambda\gamma\alpha} - \beta_{\gamma\alpha}\right) + \left(\beta_{\lambda\gamma\upsilon} - \beta_{\gamma\upsilon}\right) + \left(\beta_{\lambda\alpha} - \beta_{\alpha}\right) + \left(\beta_{\lambda\upsilon} - \beta_{\upsilon}\right) + \left(\beta_{\lambda\gamma} - \beta_{\gamma}\right) + \left(\beta_{\lambda} - \beta_{(1)}\right)\right]$$

where Δ_1 and Δ_2 are two special pairs of these 8 in total.

Put the equation of Δ_3 a bit further, there is a trick[v] to determine the subscripts in such "2^k factorial" framework. Here, for me, $k = 4$ because there are four sets of parameters, each with two statuses. To determine the signs of the parameters, rely on

$$(\lambda \pm 1)(\gamma \pm 1)(\upsilon \pm 1)(\alpha \pm 1)$$

[v] Credit to lectures of *Data Science 470 – Introduction to the Design of Experiments* by Johann Gagnon-Bartsch in Fall 2024, University of Michigan.

and, for the parameters of interest, make it -1 in the parentheses, while make it $+1$ for others.

In our example, we wanted to test λ. So, we have

$$(\lambda - 1)(\gamma + 1)(\upsilon + 1)(\alpha + 1)$$

$$= \lambda\gamma\upsilon\alpha + \lambda\gamma\upsilon + \lambda\gamma\alpha + \lambda\gamma + \lambda\upsilon\alpha + \lambda\upsilon + \lambda\alpha + \lambda - \gamma\upsilon\alpha - \gamma\upsilon - \gamma\alpha - \gamma - \upsilon\alpha$$

$$- \upsilon - \alpha - (1)$$

In addition, after summing every pairs up, we need to multiply

$$\frac{1}{n \cdot 2^{k-1}}$$

to the sum ("contrast"), where n is the number of repetitions and k is the number of factors. In our example, I repeat the simulation once ($n = 1$) and there are four sets of parameters ($k = 4$), so

$$\frac{1}{1 \times 2^{4-1}} = \frac{1}{8}$$

Put together, this trick gives us the same terms (in terms of subscripts and the signs), though in a different order, as our equation for Δ_3. This is convenient for us to write down calculations to test interactions in the next step.

In Table 3, there are three ways to test the existence of interactions between any pairs of the four educational processes. Again, the first two ways are special. Suppose we want to test the interaction between origin-education association (λ) and shares of education groups (α).

For Column 1 in Table 3, we look at

$$(\beta_{\lambda\gamma\upsilon\alpha} - \beta_{\gamma\upsilon\alpha}) - (\beta_{\lambda\gamma\upsilon} - \beta_{\gamma\upsilon})$$

to determine if, controlling the other two processes as changed (γ and υ present in the subscripts), whether the influence of λ given α is present (first parentheses) is statistically significantly different from the influence of λ given α is *not* present (second parentheses).

For Column 2 in Table 3, we look at

$$(\beta_{\lambda\alpha} - \beta_{\alpha}) - (\beta_{\lambda} - \beta_{(1)})$$

where we control the other two processes as unchanged (γ and υ not present in the subscripts).

Lastly, Column 3 in Table 3 takes all 16 counterfactual scenarios into account. Our trick can help us, and we need to make it -1 in parentheses for both λ and α,

$$\begin{aligned}(\lambda - 1)(\gamma + 1)(\upsilon + 1)(\alpha - 1)& \\ &= \lambda\gamma\upsilon\alpha - \lambda\gamma\upsilon + \lambda\gamma\alpha - \lambda\gamma + \lambda\upsilon\alpha - \lambda\upsilon + \lambda\alpha - \lambda - \gamma\upsilon\alpha + \gamma\upsilon - \gamma\alpha + \gamma - \upsilon\alpha \\ &\quad + \upsilon - \alpha + 1\end{aligned}$$

to determine the signs of the subscripts. We also need to multiply $\frac{1}{8}$ to the tentative result (contrast). Note that it only tests whether there are interactions between these two educational processes. It *does not* determine the role of, say, changes in the origin-education association and the shares of education groups together in persistence trends.

Mathematical Decomposition

The Equation

Recall that our key variable of interest is the β, the slope coefficient from a linear regression that is estimated with ordinary least square (though the survey weights are taken into account). In our case, by definition,

$$\beta = \frac{\sum_{i=1}\left(O_i^p - \mu^p\right)(O_i^a - \mu^a)}{\sum_{i=1}\left(O_i^p - \mu^p\right)^2} = \frac{\frac{\sum_{i=1}\left(O_i^p - \mu^p\right)(O_i^a - \mu^a)}{N-1}}{\frac{\sum_{i=1}\left(O_i^p - \mu^p\right)^2}{N-1}} = \frac{Cov(O^p, O^a)}{Var(O^p)}$$

where O^p is parental occupation rank and O^a is adult occupation rank. μ^p and μ^a are the population mean (grand mean) of parental and adult occupation ranks, respectively.

Next, let g denotes education group. Further, let π_g denotes the probability of group g; μ_g^p and μ_g^a are the group mean of parental and adult occupation ranks of education group g, respectively. Based on the Law of Total Covariance,

$$\begin{aligned} Cov(O^p, O^a) &= E[Cov(O^p, O^a \mid g)] + Cov[E(O^p \mid g), E(O^a \mid g)] \\ &= Cov[E(O^p \mid g), E(O^a \mid g)] + E[Cov(O^p, O^a \mid g)] \\ &= Cov\left(\mu_g^p, \mu_g^a\right) + E\left[Cov\left(O_g^p, O_g^a\right)\right] \\ &= E\left\{\left[\mu_g^p - E\left(\mu_g^p\right)\right]\left[\mu_g^a - E\left(\mu_g^a\right)\right]\right\} + \left[\sum_g \pi_g \cdot Cov\left(O_g^p, O_g^a\right)\right] \end{aligned}$$

Further, note that

$$E\left(\mu_g^p\right) = \mu^p$$

because

$$\mu_g^p = \frac{1}{n_g}\sum_{i=1}^{n_g} O_{gi}^p$$

where n_g is the sample size of education group g and

$$E\left(\mu_g^p\right) = E\left(\frac{1}{n_g}\sum_{i=1}^{n_g} O_{gi}^p\right) = \frac{1}{n_g}\sum_{i=1}^{n_g} E\left(O_{gi}^p\right) = \frac{1}{n_g}\sum_{i=1}^{n_g} \mu^p = \mu^p$$

Similarly,

$$E\left(\mu_g^a\right) = \mu^a$$

Therefore,

$$\begin{aligned} Cov(O^p, O^a) &= E\left\{\left[\mu_g^p - E\left(\mu_g^p\right)\right]\left[\mu_g^a - E\left(\mu_g^a\right)\right]\right\} + \left[\sum_g \pi_g \cdot Cov\left(O_g^p, O_g^a\right)\right] \\ &= E\left[\left(\mu_g^p - \mu^p\right)\left(\mu_g^a - \mu^a\right)\right] + \left[\sum_g \pi_g \cdot Cov\left(O_g^p, O_g^a\right)\right] \\ &= \left[\sum_g \pi_g \cdot \left(\mu_g^p - \mu^p\right)\left(\mu_g^a - \mu^a\right)\right] + \left[\sum_g \pi_g \cdot Cov\left(O_g^p, O_g^a\right)\right] \\ &= \sum_g \pi_g \cdot \left[\left(\mu_g^p - \mu^p\right)\left(\mu_g^a - \mu^a\right) + Cov\left(O_g^p, O_g^a\right)\right] \end{aligned}$$

Next, if we represent $Var(O^p)$ with $(\sigma^p)^2$, we will have

$$\begin{aligned}
\beta = \frac{Cov(O^p, O^a)}{Var(O^p)} &= \frac{\sum_g \pi_g \cdot \left[\left(\mu_g^p - \mu^p\right)\left(\mu_g^a - \mu^a\right) + Cov\left(O_g^p, O_g^a\right)\right]}{(\sigma^p)^2} \\
&= \frac{1}{(\sigma^p)^2} \sum_g \pi_g \cdot \left[\left(\mu_g^p - \mu^p\right)\left(\mu_g^a - \mu^a\right) + Cov\left(O_g^p, O_g^a\right)\right] \\
&= \sum_g \pi_g \cdot \frac{1}{(\sigma^p)^2} \cdot \left[\left(\mu_g^p - \mu^p\right)\left(\mu_g^a - \mu^a\right) + Cov\left(O_g^p, O_g^a\right)\right] \\
&= \sum_g \pi_g \cdot \left[\frac{\left(\mu_g^p - \mu^p\right)\left(\mu_g^a - \mu^a\right)}{(\sigma^p)^2} + Cov\left(O_g^p, O_g^a\right) \cdot \frac{1}{(\sigma^p)^2}\right] \\
&= \sum_g \pi_g \cdot \left[\frac{\left(\mu_g^p - \mu^p\right)\left(\mu_g^a - \mu^a\right)}{(\sigma^p)^2} + \frac{Cov\left(O_g^p, O_g^a\right)}{\left(\sigma_g^p\right)^2} \cdot \frac{\left(\sigma_g^p\right)^2}{(\sigma^p)^2}\right] \\
&= \sum_g \pi_g \cdot \left[\frac{\left(\mu_g^p - \mu^p\right)\left(\mu_g^a - \mu^a\right)}{(\sigma^p)^2} + \frac{Cov\left(O_g^p, O_g^a\right)}{Var\left(O_g^p\right)} \cdot \frac{\left(\sigma_g^p\right)^2}{(\sigma^p)^2}\right] \\
&= \sum_g \pi_g \cdot \left[\frac{\left(\mu_g^p - \mu^p\right)\left(\mu_g^a - \mu^a\right)}{(\sigma^p)^2} + \beta_g^w \cdot \frac{\left(\sigma_g^p\right)^2}{(\sigma^p)^2}\right]
\end{aligned}$$

which is exactly the same as Equation 6. In other words, in this way, it is proved why the overall intergenerational persistence can be decomposed into three items, the between-education-group inequalities (origin-education and education-destination associations together), within-education-group gaps, and shares of education groups.

The Test

The roles of the three items are tested in Equation A7. To prove the equation, let's work it backwards, first expanding the terms and then regrouping them,

$$
\begin{aligned}
&\sum_g \pi_g^{old} \cdot \left(b_g^{young} - b_g^{old}\right) + \sum_g \pi_g^{old} \cdot \left(w_g^{young} - w_g^{old}\right) + \sum_g B_g^{young} \cdot \left(\pi_g^{young} - \pi_g^{old}\right) \\
&\quad = \sum_g \pi_g^{old} \cdot b_g^{young} - \sum_g \pi_g^{old} \cdot b_g^{old} + \sum_g \pi_g^{old} \cdot w_g^{young} - \sum_g \pi_g^{old} \cdot w_g^{old} \\
&\quad + \sum_g B_g^{young} \cdot \pi_g^{young} - \sum_g B_g^{young} \cdot \pi_g^{old} \\
&\quad = \sum_g \pi_g^{old} \cdot b_g^{young} + \sum_g \pi_g^{old} \cdot w_g^{young} - \left(\sum_g \pi_g^{old} \cdot b_g^{old} + \sum_g \pi_g^{old} \cdot w_g^{old}\right) \\
&\quad + \sum_g B_g^{young} \cdot \pi_g^{young} - \sum_g B_g^{young} \cdot \pi_g^{old} \\
&\quad = \sum_g \pi_g^{old} \cdot \left(b_g^{young} + w_g^{young}\right) - \sum_g \pi_g^{old} \cdot \left(b_g^{old} + w_g^{old}\right) + \sum_g \pi_g^{young} \cdot B_g^{young} \\
&\quad - \sum_g \pi_g^{old} \cdot B_g^{young} \\
&\quad = \sum_g \pi_g^{old} \cdot B_g^{young} - \sum_g \pi_g^{old} \cdot B_g^{old} + \sum_g \pi_g^{young} \cdot B_g^{young} - \sum_g \pi_g^{old} \cdot B_g^{young} \\
&\quad = \sum_g \pi_g^{young} \cdot B_g^{young} - \sum_g \pi_g^{old} \cdot B_g^{old} = \beta^{young} - \beta^{old}
\end{aligned}
$$

APPENDIX IV
CGSS Cleaning and Processing Information[vi]

Data Cleaning

Variable Names and Meaning

During the stage of data cleaning, I need these variables below:

`gender` is the *respondent*'s gender

`birthy` is the *respondent*'s year of birth

`foccupisco` is the ISCO code of the *father*'s job when the respondent was 14

`foccuptype` determines whether the *father*'s job above is an agricultural job

`moccupisco` is the ISCO code of the *mother*'s job when the respondent was 14

`moccuptype` determines whether the mother's job above is an agricultural job

`educoriginal` is the *respondent*'s reported highest level of education

`educstatus` is the status (e.g., "graduated/completed," "enrolled," "dropped out") of the respondent's reported highest level of education (`educoriginal`)

`occupisco` is the ISCO code of the *respondent*'s current job

`occuptype` determines whether the *respondent*'s job above is an agricultural job

`hukou_now` determines the *respondent*'s current *hukou* (household registration) status

`hukou_raw` determines the year in which the *respondent* obtained a non-agricultural (urban) *hukou* status

---

[vi] Texts in `Courier New` and Libertinus Sans represent variable names.

`weight` is the survey weight

Different waves of CGSS might record these variables under different names. Table A23 at the end of this appendix shows, in each wave of data, the raw column name for each variable listed above.

Indeterminable Highest Level of Education

I will exclude the respondents whose highest level of education and are undeterminable, with `educoriginal` and `educstatus`. Such respondents usually answered "other" or "I don't know" when they took the survey.

`educoriginal` exclusion –

Waves 2008-2021: "Other (Please Specify: _____)" marked by *14*

`educstatus` exclusion –

Wave 2008: Not applicable (question not asked)
Waves 2010-2015: No exclusion needed
Waves 2017-2021: "I don't know" marked *98*; "I refuse to answer" marked *99*

Unassigned ISCO Codes

The variables `foccuptype`, `moccuptype`, and `occuptype` determine the scenario for the father, mother, and respondent, respectively. There are two scenarios when the ISCO codes are not assigned, i.e., being `NA`:

1). Not working, such as *retired*, being *unemployed*, etc.:

There are no operations needed, and such rows will be excluded in data processing.

2). Working in an agriculture job:

Their `type` will be changed to *99*, indicating that an ISCO-88 code, which is the same for all individuals in this scenario, need to be assigned for them.

`foccuptype` and `moccuptype` codes indicating an agriculture job –

2008: *8*, *9*
20102021: *2*, *3*

`occuptype` codes indicating an agriculture job –

20082021: *2*, *3*

Hukou Status at Fourteen

Since the CGSS data do not directly provide the respondents' *hukou* status at 14, I use the year in which they convert their *hukou* to determine their *hukou* status at 14.

In general, I consider their *hukou* status at 14 to be –

rural *hukou*: if the respondents were still holding a rural *hukou* when surveyed
rural *hukou*: if the conversion to a *hukou* that is not agricultural happened when they aged 14 or later
urban *hukou*: if the conversion happened before they were 14

Data Processing

Educational Attainment

I denote the education group to which the respondents belong with educ, which is determined with both `educoriginal` and `educstatus`. The codes and the

associated meanings are as follows. Note that I reorganized the words to express the same ideas of the options rather than literally translating the questionnaire.

`educoriginal` codes and the indicated levels –

01 "I've never received any education."
02 Literacy Classes
03 Primary School
04 Middle School
05 Occupational School
06 Regular High School
07 Specialized Secondary School
08 Skilled Workers' School
09 Junior College, part-time
10 Junior College, full-time
11 Bachelor's, part-time
12 Bachelor's, full-time
13 Graduate Program

`educstatus` –

01 "I am enrolling in the program."
02 "I dropped out from the program."
03 "I haven't completed the program, but I have terminated my enrollment."
04 "I have graduated."

Specifically, the degrees from junior colleges *9* and *10* are post-secondary degrees that take a shorter period of time (usually 3 years) to complete compared to the time needed for earning a bachelor's degree (usually 4 years). So, a junior college degree sounds very similar to an associate degree in the US education system. Based on the above codes and my understanding of the education system in postsocialist China, I will

borrow part of the terminology of the Bloome et al. (2018:1227-1229) article and divide the entire sample into five education groups, from lowest to highest.

Group categories, abbreviations, and associated `educoriginal` codes –

less than high school (<HS):	*1*, *2*, *3*, *4*
high school diploma (HS):	*5*, *6*, *7*, *8*
an associate's degree (AA):	*9*, *10*
a bachelor's degree or above (BA+):	*11*, *12*, *13*

Then, the scenarios of `educstatus` –

1	The respondent is in the process of completing the program / earning the degree indicated by `educoriginal` → I will move them to a lower level of education.
2 or *3*	The respondent failed to earn the degree → I will move them to a lower level as well.
4	The respondent has graduated → The categorization above doesn't change.

By "moving lower," I mean that the respondent will be moved one level lower, such as from HS to <HS. For higher education, specifically, both AA and BA+ will be moved to HS. Putting all information together, in short, educ in my sample records the most up to date *completed* highest level of education. Taking the "moving to lower level" into account, the categorization based on `educoriginal` and conditional upon individuals' `educstatus` is as follows.

When `educstatus` is *1*, *2*, or *3* –

<HS	if `educoriginal` is *1*, *2*, *3*, *4*, *5*, *6*, *7*, *8*
HS	if `educoriginal` is *9*, *10*, *11*, *12*
AA	*N/A*
BA+	if `educoriginal` is *13*

When `educstatus` is *4* –

<HS	if `educoriginal` is *1*, *2*, *3*, *4*
HS	if `educoriginal` is *5*, *6*, *7*, *8*
AA	if `educoriginal` is *9*, *10*
BA+	if `educoriginal` is *11*, *12*, *13*

Table A23. Column Names for Relevant Variables from Raw CGSS Data by Wave

Variable Name	CGSS Wave								
	2008	2010	2011	2012	2013	2015	2017	2018	2021
gender	a1	a2	a2	a2	a2	a2	a2	a2	a2
birthy	a2	a3a	a3a	a3a	a3a	a301	a31	a31	a3_1
foccupisco	b12icodeisco	fisco88	iscodad	iscodad	iscodad	fisco88	isco08_f	isco08_f	isco08_f
foccuptype	b12f	a89d	a89d	a89d	a89d	a89d	a89d	a89d	a89d
moccupisco	b13jcodeisco	misco88	iscomom	iscomom	iscomom	misco88	isco08_m	isco08_m	isco08_m
moccuptype	b13f	a90d	a90d	a90d	a90d	a90d	a90d	a90d	a90d
educoriginal	a3a	a7a	a7a	a7a	a7a	a7a	a7a	a7a	a7a
educstatus	a3b	a7b	a7b	a7b	a7b	a7b	a7b	a7b	a7b
occupisco	c8codeisco	risco881	iscorp1	iscorp1	iscorp1	a59disco88	isco08_a59d	isco08_a59d	isco08_a59d
occuptype	c2	a58	a58	a58	a58	a58	a58	a58	a58
hukou_now	a14a	a18	a18	a18	a18	a18	a18	a18	a18
hukou_raw	a14b	a19	a19	a19	a19	a19	a19 *and* a19a	a19 *and* a19a	a19 *and* a19a
weight	weight	weight	weight	weight	weight	weight	weight	weight	weight

Note: For CGSS 2018-2021, `a19` and `a19a` are two different columns both needed for `hukou_raw`.

APPENDIX V
Some R Codes for Analyses[vii]

The `sample` was the dataset from which my analyses began. It contained `wave`, `cohort` (either 1964-1979 or 1980-1995), `poccupisei` (ISEI score of parental job), `gender`, `educ`, `birthy`, `age`, `occupisei` (ISEI score of adult job), `hukou_status`, and `weight`.

Percentile Rank

I calculated the percentile rank with the following written function. Note that it processed cohort-specific datasets. The ranked samples would be used for further analyses.

```
rank.mediator = function(dataset) {
  mediator_pisei <-
    dataset |>
    group_by(poccupisei) |> summarise(n_pisei = sum(weight)) |> ungroup() |>
    mutate(n_pisei_below = 0, n_cohort = sum(n_pisei))

  mediator_pisei$n_pisei_below <-
    c(0, cumsum(mediator_pisei$n_pisei)[-nrow(mediator_pisei)])

  mediator_aisei <-
    dataset |>
    group_by(occupisei) |> summarise(n_isei = sum(weight)) |> mutate(n_isei_below = 0)

  mediator_aisei$n_isei_below <-
```

[vii] Texts in `Courier New` and Libertinus Sans represent R codes.

```
    c(0, cumsum(mediator_aisei$n_isei)[-nrow(mediator_aisei)])

  data <-
    dataset |>
    left_join(mediator_pisei, join_by("poccupisei" == "poccupisei")) |>
    left_join(mediator_aisei, join_by("occupisei" == "occupisei")) |>
    mutate(
      prank = (n_pisei_below / n_cohort) + (n_pisei / n_cohort / 2),      # parent rank
      arank = (n_isei_below / n_cohort) + (n_isei / n_cohort / 2),        # adult rank
      prange_l = (n_pisei_below / n_cohort),       # parent rank range, lower boundary
      arange_l = (n_isei_below / n_cohort)         # adult rank range, lower boundary
    ) |>
    select(
      wave, gender, cohort, birthy, age,
      pisei = poccupisei, prank, educ, aisei = occupisei, arank,
      prange_l, arange_l, hukou_status, weight
    )

  return(data)
}
```

Model Simulations

Parameter Estimation

The generalized ordered logit model (Equation 2) used the `gologit2` function in Stata. Some online resources to which I referred include:

https://www3.nd.edu/~rwilliam/gologit2/index.html
https://www3.nd.edu/~rwilliam/gologit2/tsfaq.html
https://ignacioriveros1.github.io/r/2020/03/22/r_and_stata.html

```
model.gologit2.stata = function(dataset) {
```

```
  data <-
    dataset |>
    mutate(mu.p = weighted.mean(prank, w = weight)) |> mutate(centering = prank - mu.p)

  stata("gologit2 educ centering [weight = weight]", data.in = data)
}
```

For convenience, I also used the `vglm` function in R, with the option `family = cumulative(parallel = FALSE)`, to calculate the parameters. Since the direction of probabilities they calculate was different, I had the coefficients from the `vglm` function multiply -1.

```
model.gologit2 = function(dataset) {
  data <-
    dataset |>
    mutate(mu.p = weighted.mean(prank, w = weight)) |> mutate(centering = prank - mu.p)
  coefficients <-
    vglm(educ ~ centering, family = cumulative(parallel = FALSE), weight = weight, data = data) |>
    coefvlm(matrix.out = TRUE) |> t()

  rownames(coefficients) <-
    c("<HS", "HS", "AA")

  colnames(coefficients) <-
    c("alpha", "lambda")

  coefficients <-
    coefficients * -1

  return(coefficients)
}
```

The normal linear model used the written function below.

```
model.normlm = function(dataset) {
  data <-
    dataset |>
    mutate(mu.p = weighted.mean(prank, w = weight)) |> mutate(centering = prank - mu.p)

  lm.bHS <- lm(aisei ~ centering, weights = weight, data = data[data$educ == "<HS",])
  lm.HS  <- lm(aisei ~ centering, weights = weight, data = data[data$educ == "HS",])
  lm.AA  <- lm(aisei ~ centering, weights = weight, data = data[data$educ == "AA",])
  lm.BAa <- lm(aisei ~ centering, weights = weight, data = data[data$educ == "BA+",])

  coefficients <-
    rbind(coef(lm.bHS), coef(lm.HS), coef(lm.AA), coef(lm.BAa)) |>
    cbind(
      rbind(
        sqrt(var(lm.bHS$residuals)), sqrt(var(lm.HS$residuals)),
        sqrt(var(lm.AA$residuals)), sqrt(var(lm.BAa$residuals))
      )
    )

  rownames(coefficients) <-
    c("<HS", "HS", "AA", "BA+")

  colnames(coefficients) <-
    c("gamma", "beta", "upsilon")
  return(coefficients)
}
```

Persistence Simulation

With the estimated parameters above, I then simulated the ISEI score (I^{sa}).

```
model.simulate_arank = function(dataset, gologit2, normlm) {
  data <-
    dataset |>
    mutate(mu.p = weighted.mean(prank, w = weight)) |> mutate(centering = prank - mu.p) |>
    mutate(          # boundaries of education groups
```

```
    `<HS to HS` = 1 - (exp(gologit2[1,1] + gologit2[1,2] * centering) /
                                                (1 + exp(gologit2[1,1] + gologit2[1,2] * centering))),
    `HS to AA` = 1 - (exp(gologit2[2,1] + gologit2[2,2] * centering) /
                                                (1 + exp(gologit2[2,1] + gologit2[2,2] * centering))),
    `AA to BA+` = 1 - (exp(gologit2[3,1] + gologit2[3,2] * centering) /
                                                (1 + exp(gologit2[3,1] + gologit2[3,2] * centering)))
  ) |>
  mutate(
    educ.dice = runif(nrow(dataset), min = 0, max = 1),
    norm.inv = norminv(runif(nrow(dataset), min = 0, max = 1), mu = 0, sigma = 1)
  ) |>
  mutate(
    p.educ = case_when(          # predicted educational attainment
      (educ.dice < `<HS to HS`)                                  ~ "<HS",
      (educ.dice >= `<HS to HS` & educ.dice < `HS to AA`)        ~ "HS",
      (educ.dice >= `HS to AA` & educ.dice < `AA to BA+`)        ~ "AA",
      (educ.dice >= `AA to BA+`)                                 ~ "BA+"
    )
  ) |>
  mutate(
    p.aisei = case_when(         # predicted ISEI
      p.educ == "<HS"      ~ normlm[1,1] + normlm[1,2] * centering + norm.inv * normlm[1,3],
      p.educ == "HS"       ~ normlm[2,1] + normlm[2,2] * centering + norm.inv * normlm[2,3],
      p.educ == "AA"       ~ normlm[3,1] + normlm[3,2] * centering + norm.inv * normlm[3,3],
      p.educ == "BA+"      ~ normlm[4,1] + normlm[4,2] * centering + norm.inv * normlm[4,3]
    )
  )

mediator_p.aisei <-
  data |>
  group_by(p.aisei) |> summarise(n_p.aisei = sum(weight)) |> ungroup() |>
  mutate(n_p.aisei_below = 0, n_cohort = sum(n_p.aisei))

mediator_p.aisei$n_p.aisei_below <-
  c(0, cumsum(mediator_p.aisei$n_p.aisei)[-nrow(mediator_p.aisei)])

data <-
  data |>
  left_join(mediator_p.aisei, join_by("p.aisei" == "p.aisei")) |>
```

```
    mutate(p.arank = (n_p.aisei_below / n_cohort) + (n_p.aisei / n_cohort / 2))

  return(data)
}
```

Then, with this `model.simulate_arank` function above, I simulated the rank-rank slope, and re-simulated for 21,120 times to capture the sampling uncertainty. To increasing the efficiency of codes, I used CPU parallel computing.

```
model.simulate_slope = function(dataset, gologit2, normlm, name) {
  persistence <- c(NA)

  calc = function(x) {
    data <-
      dataset |> model.simulate_arank(gologit2, normlm)

    slope <-
      lm(p.arank ~ prank, data = data, weights = weight)

    persistence[1] <-
      summary(slope)$coefficients["prank", "Estimate"]
  }

  calc <- cmpfun(calc)


  sfInit(parallel = TRUE, cpus = detectCores())

  sfLibrary(tidyverse)
  sfLibrary(qapi)

  sfExport("dataset", "gologit2", "normlm", "persistence", "name")
  sfExport("model.simulate_arank")

  persistence <-
```

```
    sfClusterApplySR(1:repeat_num, calc, name = name, perUpdate = 20, restore = FALSE) |> unlist()

  sfStop()


  return(persistence)      # a vector of 21,120 simulated persistence
}
```

Counterfactual Analysis

The counterfactual analysis involved two parts – developing different combinations of parameters and simulating persistence. There are various ways to complete the former; the latter can be fulfilled with the written functions in the previous section.

References

Asadullah, M. Niaz, and Saizi Xiao. 2020. "The Changing Pattern of Wage Returns to Education in Post-Reform China." *Structural Change and Economic Dynamics* 53:137-148.

Bai, Limin. 2006. "Graduate Unemployment: Dilemmas and Challenges in China's Move to Mass Higher Education." *The China Quarterly* 185:128-144.

Bian, Yanjie. 2001. "Guanxi Capital and Social Eating in Chinese Cities: Theoretical Models and Empirical Analyses." Pp. 275-295 in *Social Capital: Theory and Research*, edited by N. Lin, K. Cook and R. S. Burt. New York: Routledge.

Bian, Yanjie. 2006. "Guanxi." in *International Encyclopedia of Economic Sociology*, edited by J. Beckert and M. Zafirovski. London: Routledge.

Bian, Yanjie, and Lulu Li. 2012. "The Chinese General Social Survey (2003-8): Sample Designs and Data Evaluation." *Chinese Sociological Review* 45(1):70-97.

Bian, Yanjie, and Yang Yang. 边燕杰、杨洋. 2019. "The Sociology of *Guanxi* as the Subject Discourse of China (作为中国主体话语的关系社会学)." *The Journal of Humanities* 《人文杂志》09:23-34. in Chinese.

Blau, Peter M., and Otis Dudley Duncan. 1967. *The American Occupational Structure.* New York: Wiley.

Bloome, Deirdre, Shauna Dyer, and Xiang Zhou. 2018. "Educational Inequality, Educational Expansion, and Intergenerational Income Persistence in the United States." *American Sociological Review* 83(6):1215-1253.

Breen, Richard. 2004. *Social Mobility in Europe*. Oxford, UK: Oxford University Press.

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

Cai, Fang, and Meiyan Wang. 2010. "Growth and Structural Changes in Employment in Transition China." *Journal of Comparative Economics* 38(1):71-81.

Chan, Kam Wing, and Li Zhang. 1999. "The *Hukou* System and Rural-Urban Migration in China: Processes and Changes." *The China Quarterly* 160:818-855.

Chan, Wing Kit, and Jiayu Zhang. 2021. "Can University Qualification Promote Social Mobility? A Review of Higher Education Expansion and Graduate Employment in China." *International Journal of Educational Development* 84:102423.

Chen, Jiaying, and Albert Park. 2021. "School Entry Age and Educational Attainment in Developing Countries: Evidence from China's Compulsory Education Law." *Journal of Comparative Economics* 49(3):715-732.

Chen, Jing, and Rui Li. 2023. "Pay for Elite Private Schools or Pay for Higher Housing Prices? Evidence from an Exogenous Policy Shock." *Journal of Housing Economics* 60:101934.

Chen, Lupin, Dani Salleh, and Suzylah Binti Sohaimi. 2024. “Exploration of Factors on School District Housing in China.” *Multidisciplinary Reviews* 7(12):2024279.

Chen, Meng. 2013. “Intergenerational Mobility in Contemporary China.” *Chinese Sociological Review* 45(4):29-53.

Cheng, Cheng. 程诚. 2024. “Social Capital and Social Stratification in China (社会资本与中国社会分层).” Pp. 327-354 in *Handbook on Social Stratification and Mobility in Contemporary China*《当代中国社会分层与流动研究手册》, edited by L. Li and B. Zhu. 李路路、朱斌. Beijing, China: China Renmin University Press. 中国人民大学出版社. in Chinese.

Department of Degree Management and Graduate Education. 2008a, “List of ‘Project 211’ Universities (‘211 工程’学校名单).” Ministry of Education of the People’s Republic of China. 中华人民共和国教育部. Retrieved from: http://www.moe.gov.cn/srcsite/A22/s7065/200512/t20051223_82762.html

Department of Degree Management and Graduate Education. 2008b, “List of ‘Project 985’ Universities (‘985 工程’学校名单).” Ministry of Education of the People’s Republic of China. 中华人民共和国教育部. Retrieved from: http://www.moe.gov.cn/srcsite/A22/s7065/200612/t20061206_128833.html

Ding, Yanqing, Yinduo Wu, Jin Yang, and Xiaoyang Ye. 2021. “The Elite Exclusion: Stratified Access and Production during the Chinese Higher Education Expansion.” *Higher Education* 82(2):323-347.

Du, Sijia, and Yaojun Li. 2024. “Unequal Parenting in China: A Study of Socio-Cultural and Political Effects.” *The Sociological Review* 72(3):691-713.

Duan, Yide, Haotian Zhang, Wenfu Wang, and Xiaoyan Ao. 2022. “The Effects of China’s Higher Education Expansion on Urban and Rural Intergenerational Mobility.” *China Economic Review* 73:101793.

Featherman, David L., and Robert M. Hauser. 1978. *Opportunity and Change*. Academic Press.

Fei, Xiaotong. 1992. *From the Soil, the Foundations of Chinese Society: A Translation of Fei Xiaotong’s Xiangtu Zhongguo*. Berkeley, CA: University of California Press.

Feng, Siyuan. 2021. “The Evolution of Shadow Education in China: From Emergence to Capitalisation.” *Hungarian Educational Research Journal* 11(2):89-100.

Fischer, Claude S., and Michael Hout. 2006. *Century of Difference: How America Changed in the Last One Hundred Years*. New York: Russell Sage Foundation.

Ganzeboom, Harry B. G., Paul M. De Graaf, and Donald J. Treiman. 1992. “A Standard International Socio-Economic Index of Occupational Status.” *Social Science Research* 21(1):1-56.

Ganzeboom, Harry B. G., and Donald J. Treiman. 1996. “Internationally Comparable Measures of Occupational Status for the 1988 International Standard Classification of Occupations.” *Social Science Research* 25(3):201-239.

Hao, Lingxin, Alfred Hu, and Jamie Lo. 2014. "Two Aspects of the Rural-Urban Divide and Educational Stratification in China: A Trajectory Analysis." *Comparative Education Review* 58(3):509-536.

Hout, Michael. 1988. "More Universalism, Less Structural Mobility: The American Occupational Structure in the 1980s." *American Journal of Sociology* 93(6):1358-1400.

Hout, Michael, and Thomas A. DiPrete. 2006. "What We Have Learned: RC28's Contributions to Knowledge About Social Stratification." *Research in Social Stratification and Mobility* 24(1):1-20.

Hu, Anning. 2013. "Proliferation of Educational Credentials, Changing Economic Returns, and Rising Occupational Education Requirements: Evidence in Urban China from 2003 to 2008." *International Sociology* 28(4):448-466.

Hu, Anning, and Jacob Hibel. 2014. "Changes in College Attainment and the Economic Returns to a College Degree in Urban China, 2003–2010: Implications for Social Equality." *Social Science Research* 44:173-186.

Hu, Anning, and Jacob Hibel. 2015. "Increasing Heterogeneity in the Economic Returns to Higher Education in Urban China." *The Social Science Journal* 52(3):322-330.

Hu, Yongmei, Wenfeng Fan, and Weili Ding. 胡咏梅、范文凤、丁维莉. 2015. "Does Shadow Education Aggravate Inequality of Educational Outcomes: An Empirical Study on PISA 2012 Shanghai Data (影子教育是否扩大教育结果的不均等——基

于 PISA 2012 上海数据的经验研究)." *Peking University Education Review*《北京大学教育评论》13(3):29-46. in Chinese.

Huang, Youqin, and Si-ming Li. 2014. *Housing Inequality in Chinese Cities*. 1st ed. Abingdon, VA: Routledge.

Huo, Yujia, and Jane Golley. 2022. "Intergenerational Education Transmission in China: The Gender Dimension." *China Economic Review* 71:101710.

Kang, Lili, Fei Peng, and Yu Zhu. 2021. "Returns to Higher Education Subjects and Tiers in China: Evidence from the China Family Panel Studies." *Studies in Higher Education* 46(8):1682-1695.

Knight, John, and Linda Yueh. 2008. "The Role of Social Capital in the Labour Market in China." *Economics of Transition* 16(3):389-414.

Li, Chunling. 2014. "Educational Experiences and Inequality of Opportunity among the Post-80s Generation - with Comments on the 'the Silent Revolution'." *Social Sciences in China* 4:66-77.

Li, Fengliang, Fan Wu, and Yaqin Meng. 2022. "Changing Trends of Rates of Return: A Study from the Higher Education Market in China." *Journal of Marketing for Higher Education*:1-16.

Li, Haizheng. 2003. "Economic Transition and Returns to Education in China." *Economics of Education Review* 22(3):317-328.

Li, Jun. 2013. “Job Mobility in Postreform Urban China.” *Chinese Sociological Review* 45(4):81-109.

Li, Jun, Yanfeng Gu, and Chuncen Zhang. 2015. “*Hukou*-Based Stratification in Urban China’s Segmented Economy.” *Chinese Sociological Review* 47(2):154-176.

Li, Lixu. 2004. “China’s Higher Education Reform 1998–2003: A Summary.” *Asia Pacific Education Review* 5(1):14-22.

Li, Tao, and Juyan Zhang. 2010. “What Determines Employment Opportunity for College Graduates in China after Higher Education Reform?”. *China Economic Review* 21(1):38-50.

Lin, Nan. 2017. “Building a Network Theory of Social Capital.” *Social capital*:3-28.

Liu, Limin, Wolfgang Wagner, Bettina Sonnenberg, Xiwei Wu, and Ulrich Trautwein. 2014. “Independent Freshman Admission and Educational Inequality in the Access to Elite Higher Education.” *Chinese Sociological Review* 46(4):41-67.

Luo, Zhanfu, Jin Wang, and Xue Bai. 2024. “Spatial Distribution and Influencing Factors of Shadow Education Enterprises in Primary and Secondary Schools: A Case Study of Zhengzhou’s Main Urban Area.” *Journal of Infrastructure, Policy and Development* 8(6):3297.

Ministry of Education of the People’s Republic of China. 中华人民共和国教育部. 1999. *Statistical Bulletin of National Education Development in 1998* (1998 年全国教

育事业发展统计公报). in Chinese. Retrieved from: http://www.moe.gov.cn/jyb_sjzl/sjzl_fztjgb/tnull_842.html

Ministry of Education of the People's Republic of China. 中华人民共和国教育部. 2023. *Statistical Bulletin of National Education Development in 2022* (2022 年全国教育事业发展统计公报). in Chinese. Retrieved from: http://www.moe.gov.cn/jyb_sjzl/sjzl_fztjgb/202307/t20230705_1067278.html

Mok, Ka Ho, and Alfred M. Wu. 2016. "Higher Education, Changing Labour Market and Social Mobility in the Era of Massification in China." *Journal of Education and Work* 29(1):77-97.

National Bureau of Statistics. 国家统计局. 1999. *Statistical Bulletin of the People's Republic of China on National Economic and Social Development in 1998* (中华人民共和国 1998 年国民经济和社会发展统计公报). in Chinese. Retrieved from: https://www.stats.gov.cn/sj/tjgb/ndtjgb/qgndtjgb/202302/t20230206_1901942.html

National Bureau of Statistics. 国家统计局. 2023. *Statistical Bulletin of the People's Republic of China on National Economic and Social Development in 2022* (中华人民共和国 2022 年国民经济和社会发展统计公报). in Chinese. Retrieved from: https://www.stats.gov.cn/sj/zxfb/202302/t20230228_1919011.html

National Center for Education Development Research. 国家教育发展研究中心. 2001. *2001 Green Paper on Education in China: Annual Analysis Report on China's*

Education Policy 《2001 年中国教育绿皮书：中国教育政策年度分析报告》. Beijing: Educational Science Publishing House. 教育科学出版社. in Chinese.

Nee, Victor. 1989. "A Theory of Market Transition: From Redistribution to Markets in State Socialism." *American Sociological Review* 54(5):663-681.

Nee, Victor. 1996. "The Emergence of a Market Society: Changing Mechanisms of Stratification in China." *American Journal of Sociology* 101(4):908-949.

Peterson, Bercedis, and Frank E. Harrell. 1990. "Partial Proportional Odds Models for Ordinal Response Variables." *Journal of the Royal Statistical Society. Series C (Applied Statistics)* 39(2):205-217.

Qin, Kun, Ping Luo, Binbin Lu, and Zeng Lin. 2020. "Analyzing Spatial Patterns of Intergenerational Education Mobility in China." Pp. 285-301 in *Spatial Synthesis: Computational Social Science and Humanities*, edited by X. Ye and H. Lin. Cham: Springer International Publishing.

Scherger, Simone, and Mike Savage. 2010. "Cultural Transmission, Educational Attainment and Social Mobility." *The Sociological Review* 58(3):406-428.

Shi, Wentian, Xueying Mu, Wenlong Yang, and Qinchang Gui. 2024. "The Spatial Mobility Network and Influencing Factors of the Higher Education Population in China." *Science and Public Policy*.

Song, Qiongjiang, and Cheng Yong Tan. 2024. "The Relationships between Family Socioeconomic Status, Gender Inequality, and Higher Education Selection from

1962 to 2018: A Secondary Analysis of the China General Social Survey Data." *Higher Education* 88(2):729-756.

Sun, Jie. 孙洁. 2023. "Review of the Course of the Restoration of the *Gaokao* in the Late 1970s (20 世纪 70 年代末高考恢复历程回顾)." *Archives and Construction*《档案与建设》10:98-100. in Chinese.

Torche, Florencia. 2015. "Intergenerational Mobility and Gender in Mexico." *Social Forces* 94(2):563-587.

UNESCO. n.d., "Gross Enrolment Ratio." Retrieved from: https://uis.unesco.org/en/glossary-term/gross-enrolment-ratio

Wang, Peng. 王鹏. 2024. "Research on Intergenerational Mobility in China (中国的代际流动研究)." Pp. 129-158 in *Handbook on Social Stratification and Mobility in Contemporary China*《当代中国社会分层与流动研究手册》, edited by L. Li and B. Zhu. 李路路、朱斌. Beijing, China: China Renmin University Press. 中国人民大学出版社. in Chinese.

Whyte, Martin King, and William L. Parish. 1984. *Urban Life in Contemporary China.* Chicago: University of Chicago Press.

Williams, Richard. 2006. "Generalized Ordered Logit/Partial Proportional Odds Models for Ordinal Dependent Variables." *The Stata Journal* 6(1):58-82.

Williams, Richard. 2016. "Understanding and Interpreting Generalized Ordered Logit Models." *The Journal of Mathematical Sociology* 40(1):7-20.

World Bank. 2023. *School Enrollment, Tertiary (% Gross)*. Retrieved from: https://databank.worldbank.org/metadataglossary/world-development-indicators/series/SE.TER.ENRR

World Bank Gender Data Portal. 2023. *School Enrollment, Tertiary (% Gross)*. Retrieved from: https://genderdata.worldbank.org/en/indicator/se-ter-enrr

Wu, Xiaogang, and Yu Xie. 2003. "Does the Market Pay Off? Earnings Returns to Education in Urban China." *American Sociological Review* 68(3):425-442.

Wu, Xiaogang, and Donald Treiman. 2004. "The Household Registration System and Social Stratification in China: 1955–1996." *Demography* 41(2):363-384.

Wu, Xiaogang, and Donald Treiman. 2007. "Inequality and Equality under Chinese Socialism: The *Hukou* System and Intergenerational Occupational Mobility." *American Journal of Sociology* 113(2):415-445.

Wu, Xiaogang. 2010. "Economic Transition, School Expansion and Educational Inequality in China, 1990–2000." *Research in Social Stratification and Mobility* 28(1):91-108.

Wu, Xiaogang. 2017. "Higher Education, Elite Formation and Social Stratification in Contemporary China: Preliminary Findings from the Beijing College Students Panel Survey." *Chinese Journal of Sociology* 3(1):3-31.

Wu, Xiaogang, and Bingdao Zheng. 2018. “Household Registration, Urban Status Attainment, and Social Stratification in China.” *Research in Social Stratification and Mobility* 53:40-49.

Wu, Xiaogang. 2019. “Inequality and Social Stratification in Postsocialist China.” *Annual Review of Sociology* 45(1):363-382.

Wu, Xiaogang, Zhonglu Li, and Nan Wang. 2019. “Independent Freshman Admission Program (IFAP) in China’s Higher Education: Evidence from Three National Elite Universities in Beijing.” *Chinese Sociological Review* 51(1):1-28.

Wu, Yuxiao. 2011. “Labor Market Segmentation, Job Mobility and the Two-Track Model of Chinese Urban Workers’ Acquisition of Economic Status.” *Social Sciences in China* 32(3):74-86.

Wu, Yuxiao. 吴愈晓. 2013. “The Keypoint School System, Tracking, and Educational Stratification in China, 1978-2008 (教育分流体制与中国的教育分层：1978–2008).” *Sociological Studies*《社会学研究》28(4):179-202. in Chinese.

Wu, Yuxiao, and Sijia Du. 吴愈晓、杜思佳. 2018. “The Development of Higher Education over 40 Years since China’s Reform and Opening-Up (改革开放四十年来的中国高等教育发展).” *Journal of Social Development*《社会发展研究》5:1-21. in Chinese.

Xie, Hongrong. 谢红荣. 2017. "The Influence of Social Economic Status on Academic Achievement (家庭社会经济地位对学业成绩的影响)." Master's Dissertation, Normal College, Shihezi University. 石河子大学硕士学位论文. in Chinese.

Xie, Yu. 2011. "Evidence-Based Research on China: A Historical Imperative." *Chinese Sociological Review* 44(1):14-25.

Xie, Yu. 谢宇. 2012. *Sociological Methodology and Quantitative Research*《社会学方法与定量研究》(第二版). Second Edition ed. Beijing, China: Social Sciences Academic Press. 社会科学文献出版社. in Chinese.

Xie, Yu. 2014. "Gender and Family." Pp. 495-501 in *The Oxford Companion to the Economics of China*, edited by S. Fan, R. Kanbur, S.-J. Wei and X. Zhang. Oxford, United Kingdom: Oxford University Press.

Xie, Yu, and Xiang Zhou. 2014. "Income Inequality in Today's China." *Proceedings of the National Academy of Sciences* 111(19):6928-6933.

Xie, Yu, Hao Dong, Xiang Zhou, and Xi Song. 2022. "Trends in Social Mobility in Postrevolution China." *Proceedings of the National Academy of Sciences* 119(7).

Xu, Yan, Weixuan Song, and Chunhui Liu. 2018. "Social-Spatial Accessibility to Urban Educational Resources under the School District System: A Case Study of Public Primary Schools in Nanjing, China." *Sustainability* 10(7):2305.

Xue, Haiping, and Xiaohao Ding. 薛海平、丁小浩. 2009. “A Study on Additional Instruction for Students in Cities and Towns in China (中国城镇学生教育补习研究).” *Educational Research*《教育研究》1:39-46. in Chinese.

Yang, Changlin, Dingjie Zhou, Miaoling Zou, Xianliang Yang, Qifu Lai, and Feixiang Liu. 2024. “The Impact of Social Capital on Rural Residents’ Income and Its Mechanism Analysis - Based on the Intermediary Effect Test of Non-Agricultural Employment.” *Heliyon* 10(14).

Yang, Dennis Tao. 2005. “Determinants of Schooling Returns during Transition: Evidence from Chinese Cities.” *Journal of Comparative Economics* 33(2):244-264.

Yang, Po, and Rong Wang. 2020. “Central-Local Relations and Higher Education Stratification in China.” *Higher Education* 79(1):111-139.

Yang, Po. 2024. “The Winner’s Curse? Temporal and Spatial Impacts of Higher Education Expansion on Graduate Employment and Social Mobility.” *Studies in Higher Education* 49(2):286-307.

Yang, Tianyu, and Tianfang Zhang. 2024. “Social Capital Meets Guanxi: Social Networks and Income Inequality in China.” *China Economic Review* 83:102094.

Ye, Hua. 2015. “Key-Point Schools and Entry into Tertiary Education in China.” *Chinese Sociological Review* 47(2):128-153.

Yeung, Wei-Jun Jean. 2013. “Higher Education Expansion and Social Stratification in China.” *Chinese Sociological Review* 45(4):54-80.

Yu, Shuo, Ying Liu, Rui Guo, and Zhiwei Chen. 2023. "'Can We Get an Ideal Job?': The Distribution and Influencing Factors of the Occupational Stratum of Chinese College Graduates." *Sage Open* 13(4):21582440231206893.

Zhai, Xuewei. 翟学伟. 2009. "*Guanxi* or Social Capital? (是‘关系’，还是社会资本)." *Society*《社会》29(01):109-121. in Chinese.

Zhang, Lingqing, Jun Liu, Raymond Loi, Victor P. Lau, and Hang-yue Ngo. 2010. "Social Capital and Career Outcomes: A Study of Chinese Employees." *The International Journal of Human Resource Management* 21(8):1323-1336.

Zhang, Wei, and Mark Bray. 2018. "Equalising Schooling, Unequalising Private Supplementary Tutoring: Access and Tracking through Shadow Education in China." *Oxford Review of Education* 44(2):221-238.

Zhao, Yeqin. 赵晔琴. 2016. "From Graduate Job Assignment to Autonomous Employment: Individuals and the State in the Employment Relationship - Analysis of Reports in *People's Daily* from 1951 to 1999 (从毕业分配到自主择业：就业关系中的个人与国家——以 1951～1999 年《人民日报》对高校毕业分配的报道为例)." *Journal of Social Sciences*《社会科学》4:73-84. in Chinese.

Zheng, Yanqiao, Xiaoqi Zhang, and Yu Zhu. 2021. "Overeducation, Major Mismatch, and Return to Higher Education Tiers: Evidence from Novel Data Source of a Major Online Recruitment Plantform in China." *China Economic Review* 66:101584.

Zhou, Tiancheng. 2018. “Capitalization of Public School Quality into Housing Prices: An Empirical Analysis Based on School District Housing for Public Primary Schools in Shanghai.” *Chinese Studies* 7(4):286-327.

Zhou, Xiang. 2014. “Increasing Returns to Education, Changing Labor Force Structure, and the Rise of Earnings Inequality in Urban China, 1996–2010.” *Social Forces* 93(2):429-455.

Zhou, Xiang, and Yu Xie. 2019. “Market Transition, Industrialization, and Social Mobility Trends in Postrevolution China.” *American Journal of Sociology* 124(6):1810-1847.

Zhou, Xiaolu, Xixi Li, and Yaoming Gao. 2016. “Career Guidance and Counseling in Shanghai, China: 1977 to 2015.” *The Career Development Quarterly* 64(3):203-215.

Zhou, Xueguang, Nancy Brandon Tuma, and Phyllis Moen. 1997. “Institutional Change and Job-Shift Patterns in Urban China, 1949 to 1994.” *American Sociological Review* 62(3):339-365.

Zhou, Xueguang, Phyllis Moen, and Nancy Brandon Tuma. 1998. “Educational Stratification in Urban China: 1949-94.” *Sociology of Education* 71(3):199-222.

Zhuo, Ni, Chunhui Ye, and Chen Ji. 2023. “Human Capital Matters: Intergenerational Occupational Mobility in Rural China.” *Applied Economics* 55(1):1-19.