
The Harvard community has made this article openly available. Please share how this access benefits you. Your story matters

Citation

Citable link
http://nrs.harvard.edu/urn-3:HUL.InstRepos:17467182

Terms of Use
This article was downloaded from Harvard University’s DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA
Towards a General Theory of Education-based Inequality and Mobility:
Who Wins and Loses under China’s Educational Expansion, 1981-2010

A dissertation presented
by
Maocan Guo
to
The Department of Sociology
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy
in the subject of
Sociology

Harvard University
Cambridge, Massachusetts

May 2015
Towards a General Theory of Education-based Inequality and Mobility:
Who Wins and Loses under China’s Educational Expansion, 1981-2010

Abstract

My dissertation formally develops a theory of education-based inequality and mobility to integrate the existing theoretical accounts and results in the fields. The empirical puzzle I examine is why the triangle associations among social origin, educational attainment and social destination present various patterns in different societies under educational expansion. By using a variety of cross-sectional survey data from reforming China, I illustrate that class mobility strategies, structural and institutional features in the educational system and the sociopolitical institutional context are the most important dimensions to understand how educational expansion affects education-based social stratification and inequality. My analyses demonstrate that, with China’s “bottleneck” educational opportunity structure and rising educational cost under educational expansion, we observe increasing educational inequality, declining social mobility and increasing social origin differentials in the college premium in the last three decades.
TABLE OF CONTENTS

Abstract .................................................................................................................... iii
Acknowledgements ............................................................................................... vi
List of Tables .......................................................................................................... viii
List of Figures ......................................................................................................... x
Chapter 1 ............................................................................................................... 1
Chapter 2 ............................................................................................................... 22
Chapter 3 ............................................................................................................... 69
Chapter 4 ............................................................................................................. 102
Chapter 5 ............................................................................................................. 145
Chapter 6 ............................................................................................................. 165
Chapter 7 ............................................................................................................. 217
Appendix ............................................................................................................... 224
Bibliography .......................................................................................................... 226
To My Mother and Yun, the Sources of Unconditional Love.
Acknowledgements

This dissertation would not have been possible without the mentorship, guidance and support from numerous people. My gratitude is due first and foremost to my advisor and dissertation committee chair Martin K. Whyte. I have benefited enormously from the examples Marty leads in conducting research. He has pushed me to think about the “big picture” while paying attention to the details. I have benefitted not only from his insights of China and sociology in general, but also from his phenomenal kindness, patience, and continuous support and guidance during my graduate school career.

I have benefitted enormously from the support and unique perspectives of my committee members and my mentors in the Sociology department: Christopher Winship, Jason Beckfield and Filiz Garip. I am privileged to have the opportunity to learn quantitative methods from Chris for the past years. He has always encouraged me to think one layer deeper in approaching the issues. I am grateful to Jason for his sharp comments and critiques along every step of the project. I thank Filiz for being always available and for the inspiring methodological conversations and training opportunities.

I would like to express my gratitude to a number of additional faculty, staff and students from Harvard University. Particularly I wish to thank Ezra Vogel for the intellectual community he and Marty have generously created in the form of monthly China sociology dinner meeting. I also want to thank Jessica Matteson for all the non-academic support she provides as the graduate program coordinator.

I thank my family for their love and support, including my father Meitang Guo and my sister Tong Guo. I want to thank my grandmother and mother for their sincerity and kindness. I wish they could have seen the completion of my college and graduate school degrees.
Last but not least, I want to thank Yun Zhou, for her wit that’s only matched by her kindness, for her unlimited love, and unwavering support. She has listened to endless descriptions of this project, believed in me when I doubted myself, pushed me forward when I dragged my feet, and above all, shown me the great joy of sharing a life in pursuit of knowledge and intellectual fulfillment.
# List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Title</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>International Comparison on the Origin-Education-Destination</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>Relationships</td>
<td></td>
</tr>
<tr>
<td>2.1</td>
<td>A Summary of Different Theoretical Explanations in Substantive</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td>Fields</td>
<td></td>
</tr>
<tr>
<td>2.2</td>
<td>Simulation Results of Educational Expansion (N=200,000)</td>
<td>55</td>
</tr>
<tr>
<td>3.1</td>
<td>Educational Expansion in China, National Statistics, 1977-2010</td>
<td>76</td>
</tr>
<tr>
<td>3.2</td>
<td>Educational Revenues by Levels, National Statistics, 1995-2010</td>
<td>81</td>
</tr>
<tr>
<td>3.3</td>
<td>Educational Costs for Senior High School and College, National</td>
<td>84</td>
</tr>
<tr>
<td></td>
<td>Statistics, 1995-2010</td>
<td></td>
</tr>
<tr>
<td>3.4</td>
<td>Transition Rates to Senior High School by School Localities,</td>
<td>90</td>
</tr>
<tr>
<td></td>
<td>National Statistics, 1985-2003</td>
<td></td>
</tr>
<tr>
<td>3.5</td>
<td>Labor Market Composition by Education in China, National Statistics,</td>
<td>96</td>
</tr>
<tr>
<td></td>
<td>1997-2010</td>
<td></td>
</tr>
<tr>
<td>4.1</td>
<td>The Six-category EGP Schema</td>
<td>119</td>
</tr>
<tr>
<td>4.2</td>
<td>Descriptive Statistics, by Level of Transitions, the Pooled Sample,</td>
<td>120</td>
</tr>
<tr>
<td></td>
<td>1981-2010</td>
<td></td>
</tr>
<tr>
<td>4.3</td>
<td>Transition Rates by Educational Transitions and Periods, the Pooled</td>
<td>124</td>
</tr>
<tr>
<td></td>
<td>Sample, 1981-2010</td>
<td></td>
</tr>
<tr>
<td>4.4</td>
<td>Estimated Marginal Effects for Binary Logit Models on Sequential</td>
<td>127</td>
</tr>
<tr>
<td></td>
<td>School Transit, 1981-2010</td>
<td></td>
</tr>
<tr>
<td>4.5</td>
<td>Estimated Marginal Effects for Binary Logit Models on Post-</td>
<td>129</td>
</tr>
<tr>
<td></td>
<td>compulsory Transition, by Periods</td>
<td></td>
</tr>
<tr>
<td>4.6</td>
<td>Estimated Marginal Effects for Multinomial Logistic Regression on</td>
<td>133</td>
</tr>
<tr>
<td></td>
<td>Upper Second School Transitions, by Periods</td>
<td></td>
</tr>
<tr>
<td>4.7</td>
<td>Restricted Maximum Likelihood Estimation for Hierarchical Logit</td>
<td>137</td>
</tr>
<tr>
<td></td>
<td>Random Coefficient Models on College Attainment, 1981-2010</td>
<td></td>
</tr>
<tr>
<td>5.1</td>
<td>Summary Statistics for Dependent and Independent Variables</td>
<td>159</td>
</tr>
</tbody>
</table>
# List of Figures

| Figure 1.1 | The OED Relationships Predicted by the Liberal Thesis of Industrialization | 7 |
| Figure 2.1 | A Conceptual Framework of the Effects of Educational Expansion in a Given Setting | 28 |
| Figure 2.2 | A Conceptual Framework of the Effects of Educational Expansion in Varying Settings | 31 |
| Figure 2.3 | The Process of Educational Decision-Making | 36 |
| Figure 2.4 | The Direction of the Change of the Odds-ratios Give $\varepsilon$ and $\Delta w_a$ | 44 |
| Figure 2.5 | The Ability Distribution with Changing Ability Requirements | 47 |
| Figure 2.6 | An Illustrative Distribution of Newly-Created Opportunity between Groups | 48 |
| Figure 2.7 | A Summary of Theoretical Modeling and Parameters | 65 |
| Figure 3.1 | The Structure of China’s Full-time Educational System | 72 |
| Figure 3.2 | National Statistics of the Transition Rates, by Levels of Education | 78 |
| Figure 3.3 | Tuition and Fees per Student for Senior High School and College (in RMB), National Statistics, 1995-2010 | 85 |
| Figure 3.4 | The Rates of Average College Cost over GDP per capital and Household Incomes, National Statistics, 1995-2010 | 87 |
| Figure 3.5 | National Statistics of the Transition Rates to Senior High School, by School Localities, 1985-2003 | 92 |
| Figure 3.6 | The Increasing Supply of College-educated Workers, National Statistics, 1977-2010 | 95 |
| Figure 3.7 | The Changing Composition of Educational Groups for the Young Cohorts (Aged 25-29), National Statistics, 1997-2010 | 98 |
| Figure 4.1a | Trend of the Effect of Hukou Origin on Senior High School Transition | 135 |
| Figure 4.1b | Trend of the Effect of Hukou Origin on College Transition | 135 |
Figure 4.2a  Scatterplot between the Bottleneck Structure and *Hukou* Origin Effect on College Attainment, Fitted Values, over Years of Transition 139

Figure 4.2b  Fitted Slopes of Father’s ISEI on College Attainment, by Level of Constraint in the Bottleneck Structure 139

Figure 4.3  Scatterplot between College Tuition and *Hukou* Origin Effect on College Attainment, Fitted Values with 95% Confidence Intervals, over Years of Transition 140

Figure 6.1  The Percentage of Educated Workers in the Labor Market, National Statistics, 1997-2010 178

Figure 6.2  Teacher-student Ratios and Number of College, National Statistics, 1977-2010 180

Figure 6.3  Identifying the Compliers from the Observed Data 195

Figure 6.4  Rate of Yearly Schooling Return and the College Premium, Urban China, 1978-2010 201

Figure 6.5  The Changing Composition of Educational Groups, by *Hukou* Origin 202

Figure 6.6  The Distribution of the Propensity Score, by Periods of College Transition 205

Figure 6.7  MTE as a Function of Unobserved Heterogeneity (U_s): Parametric-normal approach 207
Chapter 1

Understanding the Consequences of Educational Expansion: A Theoretical and Empirical Agenda

INTRODUCTION

Educational expansion is one of the most visible, durable and consequential features of modern society. In many countries, there is a widespread tendency of increasing school enrollment across time (Shavit et al. 2007). Such a process has changed not only the overall distribution of educational opportunity (Mare 1979, 1980, 1981), but also the quantity and composition of skilled labor in the labor market (Breen 2010a; Fortin 2006; Gerber and Chang 2008; Hendricks et al. 2014; Juhn et al. 2005). As a result, it contributes to the dynamics of education-based social stratification and inequality.\(^1\)

Social scientists have long been interested in questions about the access to and the impact of education under educational expansion (Alon 2009; Breen 2010a; Breen and Jonsson 2005; Hout 2010, 2012, 2015; Mare 1981, 1991; Shavit and Westerbeek 1998; Torche 2011). Within explicit or implicit context of educational expansion, an extensive body of literature has covered a wide-range of topics including educational transition (e.g., Mare 1981, 2011), school-to-work transition (e.g., Shavit and Muller 1998), social mobility (e.g., Erikson and Goldthorpe 1992), and changing returns to education in the labor market (e.g., Carneiro and Lee 2007, 2011; Hout 2012). Although scholarly work on each of the above topics helps to illustrate a particular aspect

\(^1\) I use the terminology of “education-based social stratification and inequality” as a summary for education-related consequences in aspects including educational inequality, education-occupation (first and current) associations, and income returns to education. It is interchangeable to the term “education-based inequality and mobility” in this dissertation.
of the consequences of educational expansion, the answer to the general question about how educational expansion affects the patterns of education-based inequality and mobility is still theoretically and empirically inconclusive.

First, existing studies have usually ignored that educational expansion produces multifaceted consequences that are intrinsically connected (Hoxby and Terry 1999; Wang et al 2014; Zheng 2009). When educational expansion happens, it first affects how different social groups make use of educational opportunities. Subsequently, as students graduate from school and enter the labor market, the expansion in turn implicates how various social groups receive occupational and material returns to educational credentials. Because of the simple fact that schools provide labor forces for the labor market, how individuals access education under educational expansion is not independent of how their education is later rewarded in the labor market. In this regard, without integrating the various consequences of educational expansion, we can hardly understand the whole process of how educational expansion shapes the allocation of educational and occupational opportunities as well as the distribution of material resources in the labor market.

Secondly, comparative research has found both similarities and variations in the patterns of education-based inequality and mobility (Breen and Karlson 2013; Breen and Luijkx 2004a, 2004b; Erikson and Goldthorpe 1992, 2002, 2010; Hout 2010, 2015; Hout and DiPrete 2006). Particularly, while there is an increasing sign that industrialized (and post-industrial) countries share different education-based stratification patterns from countries such as Russia and China (Breen 2004; Breen and Jonsson 2005; Gerber 2000, 2004; Gerber and Bian 2008; Gerber and Hout 1995, 1998, 2004; Guo and Wu 2009; Wu 2010, 2011; Wu and Treiman 2007; Wu and Zhang 2010), to what extent this “empirical puzzle” holds in the context of educational
expansion and why are open to further empirical examination and theoretical explanation. Take
the case of China for example, whether its empirical patterns are indeed different from the
experience of the industrialized societies is yet to be closely investigated (Bian 2002; Wu and

Thirdly, there still lacks a coherent theory that incorporates various approaches in the
literature regarding the impacts of educational expansion. Currently, there are two branches of
theoretical explanations in the fields. On one hand, the micro-level rational choice approach tries
to theorize how students with different social backgrounds make decisions for educational
attainment and social mobility (Breen and Goldthorpe 1997, 2001; Breen and Yaish 2006;
Goldthorpe 1998, 2007, 2010). On the other hand, the macro-level institutional and structural
approaches tend to emphasize how institutional and structural changes influence social
stratification and inequality patterns (Arum 2000; Beller and Hout 2006; Breen 2010a; Breen and
Despite the importance of the micro-macro linkage in sociological studies (Coleman 1986, 1994;
Hedstrom and Swedberg 1998), thus far no dialogue between the two branches is available to
account for the varying consequences of educational expansion.

Finally, and more seriously, partially due to the missing micro-macro linkage in the
literature, current research cannot fully explain why the patterns of education-based inequality
and mobility in certain societies are different from some others’. In particular, we cannot
convincingly incorporate the “outlier patterns” observed in certain industrializing societies such
Gerber and Schaefer 2004) into the existing available theories. This is partly because current
theories mainly draw from the experience of western industrial societies (e.g., Breen et al 2009a,
2009b; Hout 2015; Hout and DiPrete 2006; Scherer et al 2007; Shavit and Blossfeld 1993; Shavit et al 2007). However, it must be recognized that as the institutional and structural changes introduced by educational expansion differ by societal settings, the impacts of educational expansion on education-based inequality and mobility may differ accordingly. In this sense, demonstrating and explaining the experience of certain industrializing countries such as China will deepen our understanding of the consequences of educational expansion.

Considering all four limitations, in this dissertation I attempt to examine the consequences of educational expansion from a dynamic perspective by integrating the existing micro- and macro-level approaches and provide general theorization of the empirical patterns. The general question I ask is: *when and how does educational expansion have different consequences for education-based inequality and mobility?* To answer this question, I will develop a general theory of educational expansion in order to understand how and why the empirical patterns of education-based inequality and mobility vary in different societies. I will then use the case of reforming China to empirically demonstrate and test the theory.

China is an especially relevant research site for this exercise. This is not only because that the case of China has rarely been included in systematic comparative education-based stratification research, but also because China has particular institutional and structural characteristics in the reform periods that are rarely found elsewhere (Deng and Treiman 1997; Guo and Wu 2009; Treiman 2012, 2013; Wu 2010; Wu and Treiman 2007; Zhou 1998; Zhou et al 1996). A comparison of reforming China with other societies thus provides a valuable chance to build up stylized facts and to formulate new theoretical explanations for the empirical variations in the literature. In this sense, my use of the Chinese case is not simply to illustrate the social stratification and inequality patterns in the reform period, which is interesting and
necessary in its own right, but more importantly, is to utilize the case as an opportunity to further develop the existing theoretical frameworks.

To achieve this purpose, for the empirical chapters, I employ a variety of Chinese survey data to specifically investigate the following questions: first, how does educational expansion affect the distribution of educational opportunities among different social groups? Secondly, what are the effects of educational expansion on the school-to-work transition? Thirdly, how does educational expansion have impacts on the linkage between social origin and destination? Finally, how does educational expansion contribute to the changing income returns to education?

Although these questions seem to span from educational attainment to education-based labor market inequality, I do not choose them randomly. First, they are among the core questions in comparative social stratification and inequality studies within which I situate my dissertation project (Ganzeboom et al 1991; Hout 2010, 2015; Hout and DiPrete 2006; Treiman and Ganzeboom 2000). Secondly, these questions, step by step, correspond to the processes of education-based social stratification and inequality, and build up important dimensions of the consequences of educational expansion. Thirdly, these topics are closely related to one another. Together, they point to a broader question: who wins and loses under China’s educational expansion? Or more generally, does educational expansion help a society to get more equality of opportunity and of condition (Breen 2010b; Jencks et al 1972; Morgan 2006)?

In the following sections, I first review the existing theoretical and empirical background and synthesize the current literature. I then briefly outline my theoretical model and formulate my main arguments and contributions. Finally, I provide an overview of the dissertation.
THEORETICAL AND EMPIRICAL BACKGROUND

In this dissertation, I cover three aspects of the consequences of educational expansion: educational transition, school-to-work transition (and/or intergenerational mobility), and income returns to education. Rather than going over each field one by one, I follow the conceptual distinction between “inequality of opportunity” and “inequality of condition” (Breen and Jonsson 2005; Demeuse 2004), and review the fields under these two themes respectively.

Inequality of Opportunity: The Origin-Education-Destination Triangle

Educational expansion is a consequential macro-level process that shifts up the mean level of educational attainment in a society. When it comes to inequality of opportunity, what interests sociologists the most is usually the triangle associations among social origin, educational attainment, and social destination under educational expansion (Breen 2010b; Hout 2015; Sobel et al 1998).

The liberal thesis of industrialization suggests that technological and economic advance tends to promote increasing importance of education, or meritocracy, as the mechanism of status transmission (Blau and Duncan 1967; Erikson et al 2010; Hout 2010; Grusky 1983; Grusky and Hauser 1984; Treiman 1970). As a consequence, there tends to be greater equality of opportunity with respect to both educational and occupational attainment as nations industrialize. In other words, as summarized in Figure 1.1, the association between individuals’ social origin and the level of their educational attainment tends to weaken; the association between individuals’ educational attainment and their class positions tends to strengthen; and finally, controlling for education, the association between social origin and
destination tends to decline (Treiman and Yip 1989; Whelan 2002). These trends, as the thesis continues to argue, should be convergent in temporal and comparative perspectives.

Empirical results, however, do not support such claims. On the association between origin and education, for example, scholars have reported a general pattern of “persistent inequality” of educational attainment in most industrial societies (Shavit and Blossfeld 1993). With few exceptions, such as Sweden, the Netherlands and Germany where patterns of declining inequality were found (Erikson and Jonsson 1996; Gamoran et al 1998; Jonsson et al 1996; Kesler 2003; Muller and Karle 1993), the effect of social origin on educational attainment is highly stable across time. Although more recently, there are increasing evidences of declining educational inequality in Europe (Breen et al 2009a, 2009b), according to Shavit and Blossfeld’s (2007) reassessment of the literature, even when such reduction in educational inequality was reported, the size of the reduction is moderate. Therefore, Shavit and Blossfeld (2007) conclude that the persistent inequality agreement is probably wrong in its strong version, but a weaker version can still stand if allowing for declining inequalities at lower educational levels and for

Figure 1.1: The OED Relationships Predicted by the Liberal Thesis of Industrialization

Educational Attainment

<table>
<thead>
<tr>
<th>Origin (Father’s class)</th>
<th>Destination (Son’s class)</th>
</tr>
</thead>
</table>

- Weakening association
- Strengthening association
declining inequalities in the middle of the 20th century. Furthermore, evidences of increasing, rather than persistent or decreasing, educational inequality at certain levels have been found in other places, such as Russia and China (Gerber 2000; Gerber and Hout 1995; Guo and Wu 2009; Wu and Zhang 2010). It is thus clear that the effect of social origin on educational attainment does not share a convergent pattern, and its size tends not to weaken over time in all places.

On the association between education and destination, the patterns become even more complex. One reason is that educational systems differ significantly across countries (Kerckhoff 1995a, 2001). In particular, the institutional context of education varies by the ways in which educational qualifications are produced and are subsequently viewed and used by employers (Konig and Muller 1986). The linkage between education and jobs is therefore system-specific, dependent on how education is organized and linked to the labor market (Allmendinger 1989; Arum 2000; Arum and Shavit 1995). On this front, empirical results show that there are apparent institutional variations in the associations between education and social destination (Rosenbaum and Kariya 1989; Rosenbaum et al 1990). The effects of education in the occupational attainment process are therefore “systematically conditioned by the respective institutional contexts” (Shavit and Muller 1998: 36). Large differences are found not only in the magnitude but also in the shape of education’s effects on class destinations (Heath and Cheung 1998; Muller et al 1998). In a recent comparative study on social mobility in Europe, Breen and his colleagues (2004) have found that when controlling for class origin, the effects of education on class destination have grown weaker and varied in size in several European countries from 1970 to 2000. These findings further reject the liberal thesis of industrialization.

On the association between origin and destination, or the pattern of intergenerational mobility, again, there is no convergent trend of greater social fluidity. Instead, scholars find a
largely resistant association between origin and destination over time and across countries (Erikson and Goldthorpe 1992; Featherman et al 1975; Jones et al 1994; Lipset and Zetterberg 1959). As Erikson and Goldthorpe (1992: 367, 378) put it, there is “a high degree of temporal stability” and “broad cross-national commonality” in relative mobility rates. Even though significant deviations from their model of “core fluidity” appear, they suggest that the deviations are neither substantial nor systematic. Recent studies, on the other hand, do report more and more substantial cross-national and cross-temporal differences in the intergenerational mobility patterns (Breen and Luijx 2004a; Ganzeboom et al 1989). In Europe, for example, researchers find that there are significant variations in the levels of social fluidity and a widespread tendency of increasing social fluidity has appeared in the last several decades (Breen 2004). In Russia, on the contrary, a pattern of declining social fluidity is observed, which challenges the empirical patterns found in Europe (Gerber and Hout 2004). It is still early to conclude that the general pattern of “constant flux” is invalid despite these variations.

In a word, comparative studies on the triangle relationships among origin, education, and destination suggest that there are considerable similarities and variations, as well as stability and changes along each pathway. Table 1.1 summarizes these various patterns and trends. It shows that while there is certainly no trend of convergence and no remarkable augment in social fluidity in all three dimensions, in many industrial societies the most visible patterns are either the persistent or declining associations in the origin-education and origin-destination relations. In other words, in these societies, changes tend to be slow and in the direction of loosening associations. In other societies, such as Russia (Gerber 2000; Geber and Hout 1995, 2004) and China (Guo and Wu 2009; Wu 2010; Wu and Treiman 2007; Zhou et al 1998), however,
Table 1.1: International Comparison on the Origin-Education-Destination Relationships

<table>
<thead>
<tr>
<th>Associations</th>
<th>Patterns of Inequality</th>
<th>Illustrative Cases</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Origin-Education Association</strong></td>
<td>Persistent Inequalities</td>
<td>Germany, France, Italy, Britain, the United States, Ireland, Spain, Hungary</td>
</tr>
<tr>
<td></td>
<td>Declining Inequalities</td>
<td>The Netherlands, Sweden, Norway</td>
</tr>
<tr>
<td></td>
<td>Increasing Inequalities</td>
<td>Soviet Union/Russia, China</td>
</tr>
<tr>
<td><strong>Education-Destination Association</strong></td>
<td>Weakening linkages</td>
<td>France, Sweden, Britain, Ireland, The Netherlands</td>
</tr>
<tr>
<td></td>
<td>Persistent Linkages</td>
<td>Russia</td>
</tr>
<tr>
<td><strong>Origin-Destination Association</strong></td>
<td>Constant Flux</td>
<td>Ireland, Britain, Germany</td>
</tr>
<tr>
<td></td>
<td>Weakening linkages</td>
<td>The Netherlands, France, Sweden</td>
</tr>
<tr>
<td></td>
<td>Increasing linkages</td>
<td>Russia, China</td>
</tr>
</tbody>
</table>


strengthening associations between origin-education and origin-destination are documented.

Why this empirical puzzle holds is yet to be identified.

To interpret the various empirical patterns, one notable explanation for the observed patterns of persistent or declining educational inequality in industrial societies is the rational choice theory (Breen and Goldthorpe 1997; Breen and Yaish 2006; Breen et al 2009a, 2009b), which takes a micro-level perspective. The theory assumes that students act rationally. Individuals choose from the different educational options available based on evaluations of the costs and benefits of each option and of the perceived probability of successful outcomes. Individuals also prioritize the avoidance of downward mobility over the achievement of upward mobility (Davies et al 2002). Educational strategies, therefore, tend to reflect parents’ efforts to maximize their children’s chances of maintaining a position that is not lower than their own. As a result, when such relative risk-aversion strategies are common among classes, class
differentials in resources, average ability level and the expectation of success become important parameters in determining class-based educational inequality. For instance, in industrial societies where relative educational costs have declined over time, the trend of declining class differentials in educational resources offsets the trend of increasing class differentials in the preference for continuing education, thus leading to the persistence of educational inequality over time (Breen and Goldthorpe 1997: 294; see also Becker 2003). Recent work further contends that economic growth and educational expansion in industrial societies have reduced class differences not only in resources, but also in children’s cognitive abilities and educational aspirations (Breen et al 2009a, 2009b). As Breen et al (2009a, p. 1480) put it, “both primary and secondary effects change in ways such that declining disparities between classes in educational attainment can be expected.”

The rational choice approach is helpful in understanding the origin-education association, as well as the other associations in the origin-education-destination triangle. In their 1995 article, Ishida et al have proposed an “inclusion and exclusion” mechanism to explain how inequalities in the triangle linkages among origin-education-destination are created and recreated. They argue that class reproduction and mobility involve different social processes, which are differentially affected by education (Hout 1988), but the way in which education mediates the association between origin and destination is mostly uniform across industrialized nations. First, higher class is more likely to include people from their origin to higher education, and to be excluded from lower education. Secondly, people with higher education tend to enter desirable social positions, and to be excluded from unfavorable positions. Although it is far from clear how the “inclusion and exclusion” mechanism takes place, the thesis itself explicitly illustrates that different classes have different mobility strategies and goals.
Following this idea, Goldthorpe (1996, 1998, 2014; Erikson et al 2005) outlines a theory of social mobility which also makes the relative risk-aversion assumption. Here, different classes are assumed to have the same goal of avoiding moving downwardly in the social hierarchy. Meanwhile, different classes have different amount of resources, and their mobility strategies tend to vary. Specifically, in order to avoid downward mobility, individuals from advantageous social origins will invest as much as possible in education and exploit means of ascription when education fails, but individuals from disadvantaged origin favor strategies that aim primarily at achieving class stability or modest advancement, and invest less in education (p. 178). In other words, maximizing educational attainment is preferable for the higher class as a way for social mobility, but not for the lower class. Instead, the lower class tends to sacrifice long-term benefits for short-term payoffs from educational attainment. In this regard, the theory implies that educational aspiration for different classes should be different unless educational opportunities are sufficient and education is relatively costless.

Unlike the rational choice approach, there is another branch of explanations that emphasize macro-level institutional and structural factors. First, there is an apparent institutional explanation for the education-destination association. Comparative empirical results show that national institutions exert strong effects on the strength and patterns between the associations of education and occupation (Andersen and van de Worfhorst 2010; Shavit and Muller 1998; Wolbers 2007). Specifically, countries that are characterized by highly stratified education systems with strong vocational components, extensive tracking, and early selection into tracks tend to show a stronger relationship between education and first occupation, when compared to countries with less vocational training and stratification (Allmendinger 1989; Saar et al 2008; Scherer 2005).
Secondly, on the origin-destination association there is an increasing theoretical interest in the institutional and structural approaches. In a recent study, Breen and Jonsson (2007) have distinguished two ways that education may impact the origin-destination relationship. The first one is an equalization process in which the declining origin-education relationship can directly weaken the origin-destination association. The second way is an expansion process that when educational inequality at higher levels is smaller than at lower levels, and when educational expansion pushes more people to come to acquire higher levels of education, the increasing share of people with higher education who are also less dependent on their social origin should reduce the overall origin-destination association. In other words, educational expansion *per se* can change the composition of socially selected groups represented in higher education and thus change the pattern of social stratification. In addition, according to Breen (2010a), the process of equalization and the process of educational expansion work differently in Sweden, Germany and Britain. In Sweden and Germany, the effect of equalization and the effect of expansion reinforce each other, but in Britain, the compositional change in higher education caused by educational expansion has little or no effect on the father-son occupational associations. In this sense, macro-based institutional and structural factors influence the process of social stratification. On one hand, institutional arrangements in the educational system not only shape the process of educational expansion but also mediate the effects on social mobility. On the other hand, educational expansion changes the structural distribution of educational opportunity and the composition of people who are able to take the opportunity. Therefore, education-based social stratification and mobility are not simply the results of the different rational choices made by different social classes.
In fact, current literature now calls for paying more attention to the institutional and structural factors in the social stratification processes. As Erikson and Goldthorpe (1987: 162) point out, patterns of social fluidity can be contingent on nation-specific historical, institutional and political characteristics that reflect a large number of underlying processes. First, structural change in the distribution of educational opportunities impacts social fluidity in various ways. To be specific, this is a question of how the marginal distribution in the mobility table as a prior may affect origin-based odds ratios (Lang and Agresti 1994). Secondly, highly specific institutions in a given nation and sustained state intervention can influence social mobility (Erikson and Goldthorpe 1992: 180-181). In particular, societies vary in their institutional arrangements of the educational system. Even among industrialized societies there still exist great variations in terms of the education-occupation associations (Gangl 2002, 2004; Muller and Gangl 2003; Shavit and Muller 1998). Thirdly, political interventions in the educational reform can also change the allocations of resources and opportunities. As many studies on socialist/post-socialist countries have demonstrated, state dynamics is an important mechanism that determines people’s life chances and social mobility in these societies (Bian 2002; Gerber 2000, 2004; Gerber and Hout 1995, 2004; Raymo and Xie 2000; Zhou and Suhomlinov 2001; Zhou and Moen 2001; Zhou et al 1996, 1997).

To summarize, current theoretical frameworks have focused either on the micro-level individual-based decision-making processes for education and mobility or on the macro-level institutional and/or structural factors. There is little effort in integrating these distinct approaches in order to explain the complex empirical patterns found in the triangle relationships among origin-education-destination (as indicated in Table 1.1).
Inequality of Condition: Income Returns to Education

Aside from the origin-education-destination triangle, educational expansion also affects material returns to education in the labor market. While empirical studies find that returns to education have changed over time (Machin and Reenen 1998; Wu and Xie 2003), there is still an ongoing debate about the sources and mechanisms of changing returns to education (e.g. Brand and Xie, 2010; Juhn et al 2005). The institutional explanation argues that institutional factors such as wage-setting institutions, training system in education, and pay-setting norms in a society explain the cross national differences in changing returns to education (Card and DiNardo 2002; Wallerstein 1999). According to this perspective, wage setting institutions (such as minimum wage, Addison and Blackburn 1999; Lee 1999), collective bargaining structure (such as unionization, Kahn 2000), and welfare state regime all moderate the effect of education on earnings. Furthermore, when a society experiences the transition from redistributive economy to market economy, the development of market institutions also creates a different incentive structure that better awards education (Nee 1989; Nee and Matthews 1996).

In contrast, the structural explanation pays more attention to the changing socioeconomic structures in the economy. The “skill-biased technology change” perspective interprets the changing returns to education using the supply and demand framework (Autor et al 1998, 2008; Acemoglu 2002). International comparisons further show that changes in the skill demand differ insignificantly among industrialized countries, but changes in the supply of skilled labor through educational expansion diverge much more (Acemoglu 2003; Card and Lemieux 2001). Several recent economic studies have discussed how the increasing supply of skilled labor under educational expansion leads to a decline in the average quality of college graduates, resulting in a decrease in the returns to college (Carneiro and Lee 2009; Fortin 2006; Juhn et al 2005). In a
word, structural changes in the composition of high-skilled workers and in the composition of jobs (Acemoglu 1998; Western et al 2009; Zhou 2014) drive the changing returns to education. In the context of a market transition economy, scholars also argue that the effect of structural change based on economic growth (Walder 2002) and labor market expansion (Xie and Hannum 1996; Zhou 2014) should be separated from the generic effect of the market.

Since changing returns to education can be determined by both institutional and structural mechanisms and the direction of the change can go both ways, it is necessary to integrate the above two perspectives. On one hand, institutional changes such as market reform can bring returns to education up; on the other hand, structural changes in the compositions of skilled labor introduced by educational expansion may deflate the price of education. Unfortunately, in the discipline of sociology, there are rarely any studies that make clear the distinctions between the two sources of effects, under the context of educational expansion.

Synthesis of Fields: Limitations and Unsolved Issues

To summarize, my literature review organized by the themes of “inequality by opportunity” and “inequality by condition” has pointed out several limitations within the current stage of research:

First, as shown in Table 1.1, it is empirically puzzling that there exists a pattern of persistent/declining associations in the linkages of origin-education and origin-destination relationships in most industrialized societies, whereas in societies like Russia and China, strengthening associations have been found. It is still unclear what the explanations are.

Secondly, there are both micro- and macro-based theoretical explanations in the fields, but these approaches have never been sufficiently integrated to adequately explain the varying empirical patterns of education-based inequality and mobility in the context of educational expansion.
Thirdly, educational expansion is only one source of the changing returns to education in the labor market. Other factors such as market-oriented institutional changes may also affect how education is materially rewarded. Finally, considering all four fields together, we find that education-based social stratification is a dynamic process which takes place step by step. Social selection in educational attainment can influence the subsequent matching of individuals to occupational positions, and in turn impact the material returns to education. However, with a few exceptions (Brand and Xie 2010; Breen and Jonsson 2007), this point is still not well articulated in the literature.

**THEORETICAL FRAMEWORK**

In the previous section, I have set up one “empirical puzzle” in the fields: why do certain societies such as Russia and China have different patterns of origin-education and origin-destination associations from most industrialized societies? I have also made two theoretical points: (1) we need to view educational expansion as a dynamic process; and (2) we need to combine micro-based rational choice approach with macro-based institutional/structural approaches. In this dissertation, driven by the above limitations in the fields, I thus formulate a general theory of educational expansion to explain various empirical patterns among different societies.

Specifically, I argue that the consequence of educational expansion in terms of education-based inequality and mobility is a function of four factors: (1) social group differences in the ability and resource distribution, (2) the different proportions of students in each social group that are able to attain schooling at a certain level before the expansion, (3) the institutional and (4)
the structural changes introduced by educational expansion. A society’s sociopolitical contexts\(^2\) as well as the institutional arrangements in the educational system before educational expansion, or “the prior”, further drive cross-national variations at different time points. The first factor captures the distribution of ability and resources for different social groups when they make rational choices regarding educational attainment and social mobility. The second factor makes up the initial condition under which educational expansion occurs. The third and the fourth factors build up the institutional and structural constraints within which rational choices are being made in the context of educational expansion. The nation-specific institutional settings summarize all the macro-level preexisting societal conditions before the educational expansion, among which, the most important elements include the state regime within which educational expansion is being enforced, the preexisting institutions in the educational system (e.g. academic tracking), and certain social institutions that affect the distribution of educational opportunities (e.g. the hukou system in the Chinese society). Since my central theoretical focus is how people’s rational choices aggregately affect education-based social stratification and inequality, the framework outlined above combines explanations based on the rational choice approach, the institutional approach and the structural approach.

I formally formulate the theory and discuss the framework in details in chapter 2. Through simulation analyses (Chapter 2) and real Chinese data (empirical chapters 4 to 6), I show that in a given society, conditioning on “the prior” (when the prior is set), the effects of educational expansion are determined by the interactions among the four factors. Among different societies where “the priors” vary, how educational expansion produces social

\(^2\) Sociopolitical institutional context includes many aspects. In the political economy literature, there are many studies on welfare state regimes (Esping-Anderson, 1990), varieties of capitalism (Crouch and Streeck 1997; Hall and Soskice 2001), state types (Levy 2006), democracy (Boix 2003) and religion (McCleary and Barro 2006). In the framework here, I mainly refer to how the state affects the organizations of the society, the educational system and the class structure.
stratification and inequality patterns is a joint conditional distribution of the four factors and “the prior”. This is because the variations in “the prior” also partially drive the variations in the four factors. Overall, I argue that analyzing the interactions among the four factors given the pre-existing nation-specific institutional settings provides comprehensive interpretations of the triangle linkages among origin, education, and destination as well as of education-based labor market inequality.

OVERVIEW OF THE DISSERTATION

The dissertation is organized as follows: Chapter 2 is the theoretical chapter in which I briefly review the different existing approaches and formalize my theory. Specifically, I argue that educational expansion is a dynamic process that produces multiple connected consequences. I also argue that educational expansion affects education-based inequality and mobility step-by-step.

Chapter 3 illustrates the necessary background information of the Chinese educational expansion process. In this chapter, I demonstrate that China has both unique structural and institutional features that are rarely observed elsewhere. I also discuss how educational expansion has affected educational attainment and education-related labor market outcomes. Based on the theoretical framework in Chapter 2, I propose three general hypotheses about the consequences of educational expansion based on China’s particular institutional contexts.

Chapter 4 investigates the relationship between social origin and educational attainment under the context of educational expansion. By analyzing educational transitions at the upper secondary level and above, I find that the advantaged social groups tend to secure educational advantages in steps. When higher education was only modestly expanded, the main advantage of the advantaged social groups was in obtaining senior high school education; when college
education was commonly available, the advantaged social groups shift their previous advantages from senior high school to college educational attainment. Based on the theoretical model, I argue that the institutional features of rising educational cost and the structural factor of a “bottleneck structure” in the Chinese educational system provide the explanations of the results.

Chapter 5 looks at the effects of social origin and educational attainment on individuals’ first job attainment across different periods, and how individuals’ first job attainment is subsequently linked to their current job positions. Utilizing a set of life history data, I compare the effects of educational attainment on social destination relative to that of social origin across time. I find that the parameter of intergenerational class immobility is highest since the late 1990s, suggesting that under the rapid educational expansion starting from 1999 there is even less social mobility. Compared to their urban counterparts, rural children’s chances of using education for social upward mobility were limited the most since the late 1990s. Based on these results, I argue that China’s recent educational reforms have ironically restricted the upward mobility channel for the disadvantaged groups (rural people in particular) through college education.

After looking into the overall relationships among social origin, educational attainment, and social destination under educational expansion, Chapter 6 examines the changes in the income returns to college education across time and investigates who have benefited more or lost less from educational expansion in incomes. By using a wide range of available data, my analyses show that college premium (the income gap between college and high school graduates) is monotonically declining between 2003 and 2008. This trend is especially salient for cohorts who have been affected by the college educational expansion introduced in 1999 (i.e. individuals were born after 1980). In addition, for these particular birth cohorts, as demonstrated by
propensity score matching methods, the effects of father’s education and class background on college premium have increased, and the urban-rural (defined by *hukou* origin) odds of college premium are also increasing. It is therefore evident that it is the children from better-off families and urban areas that benefit first and the most from the college expansion in incomes.

Chapter 7 summarizes the main findings of the three empirical chapters, and discusses the implications of the results both for the Chinese case and for the theory on educational expansion in general. It also elaborates on important theoretical and empirical contributions of this dissertation. Overall, it suggests that social selections and differentiations under educational expansion are not disjoint processes; rather, previous social processes do make up certain constraints for later socioeconomic outcomes. This chapter concludes with directions for future research.
Chapter 2

Explaining the Consequences of Educational Expansion: A General Theory

INTRODUCTION

Educational expansion is a consequential process that takes place in many societies along the course of industrialization and modernization. Within the discipline of sociology, one central concern on educational expansion is whether and how educational expansion changes the general patterns of education-based inequality and mobility (Breen 2009; Breen and Jonsson 2005; Gerber and Cheung 2008; Mare 1981; Shavit and Blossfeld 1993). Answering this question involves not only understanding how people access education, but also how educational attainment is linked to the subsequent attainment of first occupation, current occupation, and material rewards in the labor market.

The existing studies on educational expansion usually consider only a singular dimension of its multi-faceted effects on education-based inequality and mobility. For each dimension, current research documents various empirical patterns across different societies (e.g., Breen and Jonsson 2005; Erikson and Goldthorpe 1992). Taken together, there is an empirical puzzle in the fields, namely: education-based social stratification patterns tend to systematically differ in industrialized countries from those in countries like Russia and China (see Table 1.1, Chapter 1). It is still difficult to assess why this empirical puzzle holds.

More difficulties arise when it comes to theorizing and explaining the various empirical patterns documented in the literature. Across the different fields, there are theoretical explanations based on either a micro-level rational choice approach (Breen and Goldthorpe 1997; 3 Along the theme of education-based inequality and mobility, this dissertation covers three substantive fields: educational transition, school-to-work transition, and changing income returns to education.)
Breen and Yaish 2007; Goldthorpe 2007) or a macro-level institutional/structural approach
(Breen and Jonsson 2007; Mayer 2001, 2002; Mueller 1980; Shavit and Muller 1998). However,
fev studies have bridged the micro-macro linkage and integrated the different theoretical
approaches so as to provide a fully convincing explanation of why there exist various empirical
patterns. In particular, current theories cannot explain the “empirical puzzle” outlined above,
specifically, why the empirical patterns in non-industrialized societies such as Russia and China
are different (Gerber 2000, 2002, 2004; Gerber and Hout 1995, 1998, 2004; Gerber and Schaefer,
2004; Wu 2010; Wu and Treiman 2007; Zhou 1998).

The empirical and theoretical limitations partially stem from under-recognizing two key
facts about educational expansion. First, educational expansion has both direct and indirect
effects on education-based inequality and mobility, and its consequences in these aspects are
closely connected to one another. On one hand, educational expansion directly influences
individuals’ educational attainment. On the other hand, how individuals get educational
credentials aggregately determines the skill composition of the newly supplied labor forces after
they graduate from schools, which subsequently affects how individuals use education to achieve
occupational positions and obtain material resources.

Secondly, educational expansion is a macro-process that introduces both structural and
institutional changes to the educational system. Structurally, it increases the availability of school
enrollments, and thus creates new slots for schooling. Institutionally, educational expansion
changes the rules of educational attainment, i.e. how the newly-created slots are assigned. Such
structural and institutional changes make up the new scope conditions under which individuals
and their families make educational decisions and sort themselves into the “slots”. Taken
together, how educational expansion impacts stratification outcomes involves both micro- and macro-level mechanisms.

Building upon these two facts, in this chapter, I theorize the processes through which educational expansion affects the allocation of schooling opportunity as well as the distribution of later labor market outcomes. Specifically, I ask two questions: first, when and how does educational expansion generate various inequality patterns for educational attainment? Secondly, how does an individual’s social selection in educational attainment (Kerckhoff 1995a) affect their subsequent occupational and income returns to education?

In the following sections, I will first briefly review the various theoretical approaches in the fields, and propose my conceptual framework of educational expansion. Next, I formulate theoretical models to explain how educational expansion produces stepwise social inequalities. Empirically, I use simulation analyses to demonstrate how educational expansion at different levels produces various inequality patterns.

THEORETICAL APPROACHES IN THE LITERATURE

The impact of educational expansion is not only on who can now access schooling, but also on how individuals’ education is later rewarded in the labor market. Across the four substantive fields covered in this dissertation, there mainly exist three approaches for theoretical explanation.

The rational choice approach sets up a micro-level foundation for understanding education-based social stratification outcomes (Boudon 2003; Goldthorpe 1998). It has been most prevalent in the field of educational transition (Breen and Goldthorpe 1997; Breen and Yaish 2007), but has also been extended to social mobility studies (Goldthorpe 2007). With the
relative risk aversion assumption (Becker 2003; Davies et al 2002; Need and de Jong 2001; van de Werfhorst et al 2007), the rational choice approach conceptualizes educational attainment as students’ rational choice to avoid downward social mobility. At the aggregated level, it seeks to explain why there is “persistent educational inequality” in many industrialized societies (Shavit and Blossfeld 1993).

In contrast, the structural and institutional approaches adopt a macro-level perspective and use macro-level factors to explain the similarities and variations in social inequality patterns across nations and time. The structural approach in the fields usually focuses on how the changing distribution and composition of certain characteristics in the population affects the outcomes of interest (e.g., Blau 1977; Kerckhoff, 1995b). In social mobility studies, Breen and Jonsson (2007) argue that the composition of workers with origin-specific education may affect the association between social origin and destination. In the field of labor market analysis, a large body of literature in economics uses the supply-demand framework to understand educational inequalities in earnings (e.g., Acemoglu 1998; Katz and Murphy 1992); in sociology, there is also an increasing interest in linking the demographic composition to earnings inequalities (Western et al 2009; Zhou 2014).

The institutional approach in the relevant fields focuses on how institutional contexts introduce variations in observed empirical patterns in various nations and at different time points. Comparative studies show that national institutions have strong effects on the strength and association between education and first occupation (Andersen and Worfhorst 2010; Shavit and Muller 1998). In addition, specific institutions and state intervention in a particular society can drive the differences in social fluidity patterns (Breen 2004; Erikson and Goldthorpe 1992). With regard to the changing returns to education, existing studies have paid attention to certain
Table 2.1: A Summary of Different Theoretical Explanations in Substantive Fields

<table>
<thead>
<tr>
<th>Fields</th>
<th>Theoretical Approaches</th>
<th>Illustrative Unsolved Issues</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Educational Transition</td>
<td>Yes</td>
<td>How to explain rising and declining educational inequalities?</td>
</tr>
<tr>
<td>2. School-to-work transition</td>
<td>True</td>
<td>How are institutions modified by educational expansion?</td>
</tr>
<tr>
<td>3. Occupational mobility</td>
<td>Yes?</td>
<td>How to explain the rising origin-destination relationship?</td>
</tr>
<tr>
<td>4. Income Returns to education</td>
<td>Yes</td>
<td>How to disentangle the institutional mechanism with the structural mechanism?</td>
</tr>
</tbody>
</table>

Note: *Goldthorpe (2007) has a theoretical outline for social mobility based on the rational choice approach, but he rarely discusses how his theoretical agenda links to empirical findings.

institutional settings such as the market institution (Nee and Matthews 1996) and the wage-setting institution (Card and DiNardo 2002) to explain the within- and between-country differences in the changing inequality structure with respect to earnings. Table 2.1 summarizes the current general availability of the three theoretical approaches in the fields. It is evident that there is a missing linkage among the micro- and macro-level approaches.

Therefore, an integration of the existing approaches is both theoretically and empirically necessary. First, the choices individuals make are within institutional and structural contexts. When educational expansion introduces both structural and institutional changes into the educational system, individuals are not simply passive actors reacting to such changes. Instead, they consciously make educational decisions and adjust their social mobility strategies (Goldthorpe 2007) under the new constraints. Aggregately, their choices of education change the composition of skilled labor in the labor market, resulting in a potential compositional effect (Breen and Jonsson 2007) on various labor market outcomes. In this sense, both the rational choice approach and the structural/institutional approach are intrinsically connected for understanding the consequences of educational expansion.

---

4 Chapter 1 provides a comprehensive review of the literature in substantive fields. More detailed literature review of each field will also be available in respective empirical chapters.
Secondly, without a synthesis of the various approaches, we can hardly explain why there are different empirical patterns in various societies. As Chapter 1 establishes, there is an empirical puzzle in the fields. That is: while there is a general pattern of persistent or declining origin-education and origin-destination associations in most industrialized societies, for some “outliers” such as Russia and China, associations in these aspects are actually increasing. Therefore, the question remains: under what conditions does education expansion lead to persistent or declining origin-education and origin-destination inequality patterns and under what conditions are these inequalities increasing? More generally, when and how does educational expansion generate differential consequences for education-based inequality and mobility? Unfortunately, no existing explanation from a single approach can provide a convincing answer to this question.

**CONCEPTUAL FRAMEWORKS OF EDUCATIONAL EXPANSION**

In this chapter, by integrating micro- and macro-level theoretical approaches (Coleman 1986), I study when and how educational expansion affects education-based inequality and mobility. To understand these processes, I propose two conceptual frameworks for the effects of educational expansion in different settings.

Figure 2.1 provides a simplified depiction of the impacts of educational expansion in a given setting (a given society at a given time). At the micro-level, following the rational choice approach (Breen and Goldthorpe 1997), I assume that individuals make rational choices for educational attainment based on the academic ability and resources they have. At the macro-level, following the structural and institutional approaches, I conceptualize the changes introduced by
Figure 2.1: A Conceptual Framework of the Effects of Educational Expansion in a Given Setting
educational expansion in the educational system as one general structural effect and one general institutional effect. Specifically, the general structural effect refers to the change in the availability of educational opportunity, or the new schooling slots open to individuals; the general institutional effect refers to the change in the selection rule of schools, which determines how the (new) schooling slots are finally distributed among the individuals.

Since post-compulsory school admission is usually associated with some forms of selection based on merit, increasing schooling opportunity means that on average schools lower their ability requirement for students. In addition, if the students have passed the ability requirement, post-compulsory education usually further requires them to possess a certain amount of resources so as to cover the educational costs associated with schooling, which are most commonly (but not limited to) in the form of tuition and fees. Therefore, with a changing resource requirement, schools can also change the rule of selection and determine who can finally seize the newly-created educational opportunities.

Under the rational choice assumption, conscious decision makers will be aware of the changes in ability and resource requirements of the schools. Decision makers will then adjust their preferences for education under the new scope conditions (the macro-micro link in Figure 2.1), and behave accordingly based on the ability and resources they possess. As a result, not only will a larger number of students go to schools after educational expansion, but their ability and resource distributions will also be different from those attending schools before the expansion. When these post-expansion students graduate and go into the labor market, at the aggregate level, again, there will be differences not only in the total number of skilled labor, but

---

5 According to human capital theory, education augments ability and training (Rosen 1977). If we assume that schools do not differ in shifting students’ ability distribution before and after educational expansion, then the students’ ability distribution after schooling will be reduced to that before schooling. This holds similarly for the resource distribution if changes in the resource distribution are the same before and after educational expansion.
also in their ability and resource distributions, as compared to the before-expansion cohorts (the micro-macro link in Figure 2.1). Consequentially, both the overall volume and the composition of educated workers will be different before and after educational expansion.

The increasing supply of educated workers who differ from earlier cohorts in terms of ability and resources changes the demographic composition of skilled labor in the labor market. Since the cohorts who attend schools before and after the expansion are only “imperfectly substitutable” (Card and Lemieux 2001) for each other, how education is rewarded in the forms of (first and current) occupational opportunities and income returns should be different across cohorts. In Figure 2.1, I conceptualize the linkage between the changing number and composition of educated workers and the occupational and material returns to education as an indirect effect of educational expansion, which can further be moderated by individual characteristics. 6

Considering that different societies have varying degrees, paces, and types of educational expansion, for comparative purposes, I propose a second conceptual framework to understand the effects of educational expansion in various settings (i.e., different societies, or different time points in a given society), which is illustrated in Figure 2.2.

Figure 2.2 is similar to Figure 2.1, but with three modifications. First, it shifts the attention from individual-level characteristics to aggregate-level group differentials. This is

---

6 The changing labor market returns to education can also be one scope condition for rational choice makers, as people may make educational choices based on the expected returns to schooling (Cameron and Heckman 1998). This means that there is a possible backdoor linkage between (expected) occupational and material returns to education and educational attainment (Hussey and Swinton. 2011; Looker and McNutt 1989). However, the (expected) changing returns to education are not solely the result of educational expansion. They are also determined by labor market demands and regulations, as well as the information and preference individuals have. To avoid the complexity of the causal links, I do not include the potential backdoor association between (expected) returns to education and educational attainment in the conceptual frameworks in this chapter.
Figure 2.2: A Conceptual Framework of the Effects of Educational Expansion in Varying Settings
because sociologists’ interest is not to explain the actions of single individuals, but rather to understand group differentials in aggregate outcomes. In the studies of educational-based social inequality and mobility, one of the key group differentials examined in the literature is social class (Erikson and Goldthorpe 1992). Therefore, to be consistent with the existing literature, I also use social class differentials in ability and resources as illustrative indicators for aggregate level group differences.  

The second modification is that I incorporate an element of nation-specific institutional contexts into the framework. It is a composite of all the nation-specific institutional settings that determines (1) the initial class differentials in ability and resources; (2) the share of each class in schools before educational expansion; and (3) the characteristics such as the extent, pace and type of educational expansion. Identifying the relevant nation-specific institutional contexts and how they impact the above three elements is helpful for understanding variations among societies, but the identification of its precise contents is theoretically unnecessary for building a parsimonious theory of educational expansion. For now, what we need to establish is that there are certain nation-specific institutional contexts that determine the initial conditions under which

---

7 Social class is only one way of social grouping. The group differential can also refer to other groupings, such as those based on race, gender, religion, and the urban-rural divide. The theoretical interest here is not in the social grouping differentials per se, but rather, in the changes in the group differentials before and after educational expansion. In this Chapter, consistent with the existing literature, I have chosen to focus on social class; however, it is only used as an illustrative point, and the theory generated here can be extended to group differentials based on other attributes in the future.

8 Among them, one of the most important institutional features is the state regime that enforces educational expansion. For example, an authoritarian regime may be more likely than a democratic regime to have a sudden policy change on education. Other important features include economic development that determines class differentials in resources, the preexisting institutions in the educational system such as academic tracking, and social institutions that systematically affect the distribution of educational opportunities such as the hukou system in Chinese society (Wu 2010; Wu and Treiman 2007).

9 If we view Figure 2.2 as a directed acyclic graph, then the element of nation-specific institutional contexts only indirectly affects individuals’ educational attainment. It therefore becomes a confounder for the linkages among the abovementioned three elements and educational attainment. Its effect on educational attainment will be blocked if we condition on it (in a given society); otherwise, its effect will be summarized by the effects of the three elements on educational attainment, conditioning upon its respective association to the three factors.
educational expansion happens (the first two elements), and the characteristics of educational expansion (the third element).

The third modification is that since the three elements of the educational expansion process are conditioned on nation-specific institutional contexts, Figure 2.2 allows them to vary across nations. Presumably, all of the three elements, i.e. (1) the initial class differentials in ability and resources, (2) the share of each class in schools before educational expansion, and (3) the characteristics of educational expansion, will be different across societies. Furthermore, the last two elements can also be different within a society across various time points. In other words, within a given society, multiple waves of educational expansion may take place and subsequently change the “initial” condition with respect to the share of each class in schools, as well as the characteristics of educational expansion wave-by-wave.

Given that the characteristics of educational expansion (the third element) vary across countries and across time points, the structural and institutional changes introduced by the expansion can also be different, subsequently further producing various degrees and types of changes in the school requirements for ability and resources. Furthermore, at the aggregate level, there exist class differentials in the overall population distributions for ability and resources. Taken together, when educational expansion produces different school requirements for ability and resources, this leads to potentially different configurations of labor forces with varying numbers and social class compositions. As a result, how educational expansion introduces changes in social class differentials in the aspects of educational attainment, education-occupation associations, and income returns to education will differ across various settings. In this sense, Figure 2.2 is a general version of Figure 2.1.
In sum, the conceptual frameworks illustrated in Figure 2.1 and 2.2 suggest that the effects of educational expansion involve both direct and indirect processes with multi-layered mechanisms. To understand why there are various patterns, we therefore need to analyze two related issues: (1) how educational expansion affects individuals’ educational choices within the changing structural and institutional contexts; and (2) how individuals’ educational choices affect the aggregate pattern of education-related labor market outcomes.

A GENERAL THEORY OF EDUCATIONAL EXPANSION

In the previous section, I contend that educational expansion influences not only the way in which individuals are socially selected into schools within certain structural and institutional contexts, but also how individuals are rewarded by their schooling choices in the labor market. In this section, building upon the theoretical frameworks illustrated in Figure 2.1 and 2.2, I formulate a formal theory of educational expansion that bridges both the micro and macro-level approaches in the existing literature.

Educational Expansion and Educational Attainment

I start with the rational choice approach. The original rational choice theory in educational transitional studies (Breen and Goldthorpe 1997) has three main elements: (1) the structure of the decision problem within which student \( i \) faces the choice of pursuing a more risky or a less risky option; (2) a threshold \( T_i \) that determines the student’s minimum acceptable level of educational attainment; and (3) the belief each student holds regarding the probability of success for each of the risky options, or the subjective belief parameter \( \pi_i \) (Breen and Yaish
2006). Under the relative risk aversion assumption, the theory defines \( T_i \) to be a social class position that is at least as “good” as the student’s social origin.

Although the rational choice theory helps to explain the micro-level educational decision-making process, it pays inadequate attention to the structural and institutional constraints within which individuals make choices. In the following sections, I modify certain features of the rational choice theory and incorporate the macro-level structural and institutional changes in order to analyze the effects that educational expansion exerts on individuals’ schooling choices.

*The Behavior Model*

To begin, consider the single educational decision tree in Figure 2.3 (reproduced from Breen and Goldthorpe 1997: 280). Assume that the educational system consists of a certain number of levels, with \( K=1, 2, \ldots, k \) indicating each post-compulsory level. After completing compulsory education, there are two terminal educational outcomes: Students can either choose to stay for further education at the \( k^{th} \) level or to leave for the labor market (\( L_{k-1} \)). If the student chooses to stay for education at the \( k^{th} \) level, there are again two possible outcomes, i.e. (1) successfully passing all \( k \) level requirements and continuing to the \( k+1^{th} \) level (\( P_k \)) or (2) failing to complete the \( k^{th} \) level and/or leaving for the labor market after the \( k \) level education is completed (\( F_k \)). Let \( L_{k-1}, P_k \) and \( F_k \) denote these three possible educational outcomes.

The rational choice theory maintains that students consider three factors when deciding whether to continue to level \( k \): the cost of remaining in school relative to the cost of leaving, captured by a resource parameter \( c \); the expected likelihood of success if choosing to continue, captured by the subjective belief parameter \( \pi_{ik} \); and the expected utility rational actors (the students and their
families) attach to each of three educational outcomes, \(L_{k-1}, P_k\) and \(F_k\). The decision rule for student \(i\) to continue to level \(k\) therefore can be summarized as:

\[
\begin{cases} 
    r_i > c \\
    \pi_{lk} > \bar{\pi}_k \\
    U(S_{lk}) > U(L_{i(k-1)})
\end{cases}
\]  \(1\)

Here, \(r_i\) denotes the family resource for student \(i\), \(\bar{\pi}_k\) denotes the minimum subjective probability of success at level \(k\), and \(U(S_{lk})\) and \(U(L_{i(k-1)})\) are the expected utility for staying at level \(k\) (including both outcomes \(P_k\) and \(F_k\)) and leaving at level \(k-1\)(outcome \(L_{k-1}\)) for student \(i\). Given the subjective belief parameter \(\pi_k\), the expected utility of staying at level \(k\) is \(U(S_{ik}) = \pi_k U(P_k) + (1-\pi_k U(F_k))\).
\(\pi_{ik}U(F_{ik})\), so the condition \(U(S_{ik}) > U(L_{i(k-1)})\) can be rewritten as: 
\[
\pi_{ik}U(P_{ik}) + (1-\pi_{ik})U(F_{ik}) > U(L_{ik}) - U(P_{ik}) - U(F_{ik})
\]
and we get: 
\[
\pi_{ik} > \frac{U(L_{ik}) - U(F_{ik})}{U(P_{ik}) - U(F_{ik})}
\]

Under the relative risk aversion assumption, students seek to avoid downward social mobility. Therefore the expected utility associated with each educational outcome is defined by the expected probability of gaining access to certain social class positions, and the threshold \(T_i\) refers to a class position that is no worse than the student’s social origin. Particularly, assume that there are three social classes: the service class (S*), the working class (W*) and the underclass (U*), and the expected probabilities for accessing each class position are denoted by the \(\alpha, \beta,\) and \(\gamma\) parameters respectively (Breen and Goldthorpe 1997). Based on the single decision-making tree in Figure 2.3, the probability of or preference for continuing to level \(k\) education is defined as 
\[
p_{sk} = \frac{\pi_{sk}\alpha + (1-\pi_{sk})\beta_1}{\pi_{sk}\alpha + (1-\pi_{sk})\beta_1 + \gamma_1}
\]
for the service class students, and 
\[
p_{wk} = \frac{\pi_{wk}(1-\pi_{wk})(\beta_1 + \beta_2)}{\pi_{wk}(1-\pi_{wk})(\beta_1 + \beta_2) + (\gamma_1 + \gamma_2)}
\]
for the working class students.\(^{10}\)

With four assumptions about the relative size of \(\alpha, \beta,\) and \(\gamma\) parameters, Breen and Goldthorpe (1997: 282) establish that for any given \(\alpha, \beta,\) and \(\gamma\) parameters with \(\forall \pi \leq 1, p_{sk} > p_{wk}\) if \(\pi_{sk} = \pi_{wk}\). Namely, with the same subjective belief parameter \(\pi\), the service class students tend to have a higher preference for continuing education as compared to the working class students.

Similarly, we have: 
\[
\pi_{sk} < \pi_{wk} \text{ if } p_{sk} = p_{wk}. \quad (2)
\]

\(^{10}\) See Breen and Goldthorpe (1997) for a detailed discussion about the respective probabilities and parameters.

\(^{11}\) Proof: \(\pi_{sk} < \pi_{wk}\) if \(P_{sk} = P_{wk}\) with \(\forall \pi \leq 1\) if an only if 
\[
\frac{\pi_{sk}\alpha + (1-\pi_{sk})\beta_1}{\pi_{sk}\alpha + (1-\pi_{sk})\beta_1 + \gamma_1} < \frac{\pi_{wk}(1-\pi_{wk})(\beta_1 + \beta_2)}{\pi_{wk}(1-\pi_{wk})(\beta_1 + \beta_2) + (\gamma_1 + \gamma_2)}.
\]
Following Breen and Goldthorpe (1997: 284), there is 
\[
\frac{\pi_{sk}\alpha + (1-\pi_{sk})\beta_1}{\gamma_1 + \gamma_2} > \frac{\pi_{wk}(1-\pi_{wk})(\beta_1 + \beta_2)}{\gamma_1 + \gamma_2},
\]
so the equation above can be rewritten as: 
\[
\frac{\pi_{wk}(1-\pi_{wk})(\beta_1 + \beta_2)}{\gamma_1 + \gamma_2} \Rightarrow (\pi_{sk} - \pi_{wk})(1 - \beta_1 - \beta_2) < 0 \Rightarrow \pi_{sk} < \pi_{wk}.
\]
In other words, when the preference for continuing education is the same for the two groups, continuing education requires a higher subjective probability of educational success ($\pi$) for the working class students than for the service class students. As a result, since one condition for continuing education is $p_k > .5$, the minimum requirement of $\pi_k$ for the S* and W* groups under the condition of $p_{sk} = p_{wk} = .5$ follows $\pi_{sk} > \pi_{wk}$.

The subjective belief parameter, $\pi_i$, can be viewed as a function of the true (demonstrated) academic ability, $a_i$, namely $\pi_i = g(a_i)$. Given the high correlation between the true academic ability and the expected probability of educational success, the minimum ability requirement of going to level $k$ schools for the S* and W* groups follows that:

$$\pi_{sk} > \pi_{wk} \Rightarrow g(a_{sk}) > g(a_{wk}) \Rightarrow a_{sk} > a_{wk}. \quad (3)$$

In words, Inequality (3) means that the minimum academic ability requirement for continuing education is higher for the working class students than for the service class students, which only holds if $a_i$ and $\pi_i$ are positively correlated. Moreover, since the academic ability requirement is a direct measure of the selectivity of the schools, I focus on both $a_i$ (ability) and $\pi_i$ (subjective belief) parameters to indicate the likelihood of educational success at level $k$. As a result, unlike the current version of the rational choice theory, I do not assume that the ability differences among social groups will be “wholly captured in differences in the subjective parameter $\pi$.\)”(Breen and Goldthorpe 1997: 286)

Therefore, I rewrite the behavior model for educational continuation in (1). That is, student $i$ in social group $j$ (j=S* and W*) will go to level $k$ schools if and only if:

$$\left\{ \begin{array}{l} \tau_{ijk} > c_k \\ \pi_{ijk} > \pi_{jk} \\ a_{ijk} > \tilde{a}_{jk} \end{array} \right. \quad (4)$$
Here, \( c_k \) denotes the resource requirement for the level \( k \) education, \( \pi_{jk} \) and \( a_j \) for \( p_{jk} = .5 \). Taken all three conditions together, the behavior model contends that students will continue to the level \( k \) education only if their true academic ability, the expected probability of academic success at level \( k \), and their available resource all exceed the respective minimum levels of requirements.

Aggregately, the behavior model implies that students from different social origins should have different average probabilities of continuing education. One main reason for such a differential is that the levels of average ability and resource differ by social classes (Boudon 1974; Breen et al 2009a). Following the existing literature (Breen and Goldthorpe 1997; Bukodi and Goldthorpe 2013; Bukodi et al 2014), I assume that the service class (\( S^* \)) and the working class (\( W^* \)) differ only in the distributions of (demonstrated) academic ability and resources. Ability is taken to be normally distributed within each group, with means \( \bar{a}_s > \bar{a}_w \) and standard deviations the same (\( \delta_{as} = \delta_{aw} \)). It captures the “primary effect” of educational attainment, and the advantaged group on average has better academic performance (Boudon 1974).\(^{12}\) Because of the high correlation between the parameters \( \pi \) and \( a \), the mean subjective probability of educational success is also lower in the working class than in the service class, i.e. \( \bar{a}_s > \bar{a}_w \) yields \( \bar{\pi}_s > \bar{\pi}_w \). Similarly, resources are taken to have a logistic distribution with mean values \( \bar{r}_s > \bar{r}_w \), and the same dispersion parameter within the two groups.

In addition, to capture the differences of class differentials in ability and resources among various societies, define:

\[
\begin{align*}
\frac{\bar{a}_w}{\bar{a}_s} &= \theta \\
\frac{\bar{r}_w}{\bar{r}_s} &= \varphi
\end{align*}
\]

\( (5) \)

\(^{12}\) The “primary effect” suggests that ability and resources are intertwined rather than being independent factors, but as Figure 2.1 indicates, they are still two conceptually different mechanisms for educational attainment.
where $0<\theta\leq 1$, and $0<\varphi\leq 1$. By definition, the smaller $\theta$ and $\varphi$ are, the larger the group differences in resources and ability are.

**Educational Transition to Level k**

Thus far, by using the rational choice approach, I have set up a behavior model for educational choices. Nonetheless, individuals’ choices are confined by the structural and institutional changes in the educational system. As conceptualized in Figure 2.1, on one hand, educational expansion provides increasing educational opportunities to students by lowering the selection criteria on academic ability. Let $a_{k,t}^0$ and $a_{k,(t+1)}^0$ be the minimum level of ability requirement to enter level k schools at time t and t+1 ($a_{k,(t+1)}^0 < a_{k,t}^0$), respectively, and let $f(x)$ denote the density function of the ability distribution in the overall population. The structural change brought by educational expansion in the form of increasing schooling opportunities is then captured by the change in $a_k^0$ from time t to time t+1, i.e. $\Delta a = \int_{a_{k,(t+1)}^0}^{a_{k,t}^0} f(x)dx$, in which $\Delta a$ is the increased number of school enrollment slots.

The institutional change, on the other hand, determines how the newly created slots are finally assigned within the population. Based on the behavior model illustrated in (4), the resource requirement is one key institutional factor that determines who can finally obtain the newly created schooling slots. Let $r_k^0$ be the minimum level of resource requirement for level k schools, and $f(y)$ be the density function of the resource distribution in the overall population. If the ability requirement is held constant from time t to time t+1, then the shifting population that can now afford level k education is: $\Delta r = \int_{r_{k,(t+1)}^0}^{r_{k,t}^0} f(y)dy - \int_{r_{k,t}^0}^{r_{k,(t+1)}^0} f(y)dy$. 

40
Together, $\Delta a$ and $\Delta r$ suggest that educational expansion has generated changes in two aspects: (1) changes in the ability requirement, which determine who are qualified for the newly created schooling opportunity, and (2) changes in the resource requirement, which determine who are finally able to afford and thereby seize the new opportunities in schooling. Given these two changes, the population change in continuing to level $k$ education under the expansion can be written as

$$
\Delta = \int f(x)f(y)dx\,dy, \quad x \in [a_{k,(t+1)}, a_{k,t}]; \quad y \in [r_{k,(t+1)}, r_{k,t}] \text{ or } [r_{k,t'}, r_{k,(t+1)}]
$$

depending on whether the resource requirement is increasing or decreasing.

Since changes in the ability and resource requirements make up the scope conditions under which individuals make educational decisions, the behavior model (4) further requires that the academic ability and resources of student $i$ exceed $a_k^0$ and $r_k^0$, namely: \[13\]

$$
\begin{cases}
\tilde{a}_{jk} \geq a_k^0 \\
\tilde{c}_k = r_k^0
\end{cases}
$$

(6).

Together with the assumption of class differentials in the ability distribution $(\tilde{a}_s, \tilde{a}_w, \delta_{as} = \delta_{aw})$, it follows that the standardized z-scores for $\tilde{a}_{jk}$ and $a_k^0$ are lower for the $S^*$ than for the $W^*$ students, i.e.: \[\frac{a_{sk} - \tilde{a}_s}{\delta_{as}} < \frac{a_{wk} - \tilde{a}_w}{\delta_{aw}} \text{ and } \frac{a_k^0 - \tilde{a}_s}{\delta_{as}} < \frac{a_k^0 - \tilde{a}_w}{\delta_{aw}}\]. In other words, the ability percentile of students qualified for level $k$ education is higher in the working class than in the service class, when only considering ability as the determinant factor.

In educational transition studies, the most commonly used measure for educational inequality is the odds-ratio (Mare 1981). \[\text{14}\] Define Odds-ratio (OR) = $\frac{p_s/(1-p_s)}{p_w/(1-p_w)}$, where $p_s$ and $p_w$

---

\[\text{13}\] For simplicity here I assume that the ability and resource requirements are uniform across various social groups. One example in this scenario is admissions determined by school entrance exams. In cases where there is an egalitarian educational policy or an affirmative action policy, the ability and/or resource requirements may differ by social groups. However, no matter how the school requirements vary by different groups, aggregately they only determine the distribution of the new educational opportunity across groups, which are captured by $\Delta s$ and $\Delta w$ in later analyses.
are the respective proportion of the service class and the working class students who choose to continue to level k education. According to (4) and (6), write
\[ p_s = pr(a_{is} > \hat{a}_s | \hat{a}_s \geq a^0_s) pr(\pi_{is} > \hat{\pi}_s) pr(r_{is} > r^0_t), \]
\[ p_w = pr(a_{iw} > \hat{a}_w | \hat{a}_w \geq a^0_w) pr(\pi_{iw} > \hat{\pi}_w) pr(r_{iw} > r^0_t) \]
for the i^{th} student in the S* and W* groups (the subscript k omitted) in time t.

Given that \( a_i \) and \( \pi_i \) are highly correlated, for simplicity, I analytically use the ability parameter \( a_i \) to summarize the requirements in \( a_i \) and \( \pi_i \) parameters. Further, let \( \hat{a}_i = a_i^0 \) so as to simplify \( p_s \) and \( p_w \) to only reflect the impacts of the structural and institutional changes under educational expansion.\(^{15}\) We now get:
\[ p_s = pr(a_{is} > a_i^0) pr(r_{is} > r^0_t), p_w = pr(a_{iw} > a_i^0) pr(r_{iw} > r^0_t). \quad (7) \]

To understand how the structural change introduced by educational expansion affects educational inequality, for now set \( \varphi = 1 \) so that there is not group difference in resources, namely
\[ pr(r_{is} > r^0) = pr(r_{iw} > r^0) \]
at time t and t+1. Let \( p_1 = pr(a_{is} > a_i^0), p_2 = pr(a_{iw} > a_i^0), \) the odds-ratio at time t is then given by
\[ OR_t = \frac{p_1/(1-p_1)}{p_2/(1-p_2)}. \]

The term \( \Delta a \) captures the increased educational opportunities due to the lowering of ability requirement from time t to t+1. Let \( \Delta s_a \) denote the proportion of students in the service class that benefitted from the new opportunities created for the level k education, and define \( \Delta w_a \) similarly for the working class (\( 0 \leq \Delta s_a < 1 - p_1, 0 \leq \Delta w_a < 1 - p_2 \)). Furthermore, let \( \varepsilon \) be the difference between \( \Delta s_a \) and \( \Delta w_a \) so that \( \Delta s_a = \Delta w_a + \varepsilon. \) By definition \( p_2 < p_1 \) and

\(^{14}\) Although there are debates on whether odds-ratio is an appropriate measure of educational inequality (Hellevik 2000; Marshall and Swift 1999), for comparative purposes, I still employ odds-ratio to directly measure the relationship between social origin and educational attainment.

\(^{15}\) Aggregately, the level of ability requirement of schools at time t (\( a_i^0 \)) is the equilibrium of the average level of minimum requirement of ability for continuing education based on the relative risk aversion assumption (\( \hat{a}_j \)).
\( p_2 + \Delta w_a < p_1 + \Delta s_a < 1 \), then \( p_2 - p_1 < \varepsilon < 1 - p_1 \). The odds ratio at time \( t+1 \) is now defined as

\[
OR_{t+1} = \frac{(p_1 + \Delta s_a)/(1 - (p_1 + \Delta s_a))}{(p_2 + \Delta w_a)/(1 - (p_2 + \Delta w_a))}.
\]

Rewrite \( OR_t \) and \( OR_{t+1} \), we get:

\[
OR_t = \frac{p_1 - p_2}{p_2 - p_1 p_2}, \quad \text{and} \quad OR_{t+1} = \frac{p_1 - p_2 - (p_1 \Delta s_a + p_2 \Delta w_a + \Delta s_a \Delta w_a) + \Delta s_a}{p_2 - p_1 - (p_1 \Delta s_a + p_2 \Delta w_a + \Delta s_a \Delta w_a) + \Delta w_a}.
\]

Since both denominators of \( OR_{t+1} \) and \( OR_t \) are positive, when calculating \( OR_{t+1} - OR_t \), it becomes apparent that the sign of \( OR_{t+1} - OR_t \) depends on the sign of the following item:

\[
(p_1 - p_2) \times [\Delta w_a^2 - (1 - p_1 - p_2 - \varepsilon) \Delta w_a + \frac{p_2 (1-p_2) \varepsilon}{p_1 - p_2}].
\]

Since \( p_1 - p_2 > 0 \), as a quadratic function of \( \Delta w_a \) Item (9) will always be greater than 0 if

\[
M = (1 - p_1 - p_2 - \varepsilon)^2 - 4 \times \frac{p_2 (1-p_2) \varepsilon}{p_1 - p_2} = \varepsilon^2 - 2 \times \frac{p_2 (1-p_2) + p_2 (1-p_2)}{p_1 - p_2} \varepsilon + (1 - p_1 - p_2)^2 < 0,
\]

which is again a quadratic function of \( \varepsilon \). Solving the inequality \( M<0 \), we get

\[
\varepsilon \in \left( \frac{\sqrt{p_2 (1-p_2) - \sqrt{p_2 (1-p_2)}}}{p_1 - p_2}, 1-p_1 \right).
\]

In words, any \( \varepsilon \) falls into this particular range will predict a higher odds-ratio at time \( t+1 \). Similarly, if \( M>0 \), \( \varepsilon \in \left( p_2 - p_1, \frac{\sqrt{p_2 (1-p_2) - \sqrt{p_2 (1-p_2)}}}{p_1 - p_2} \right) \), by solving the
Figure 2.4: The Direction of the Change of the Odds-ratios Give $\varepsilon$ and $\Delta w_\alpha$

(2.4a): $1 - p_1 - p_2 \geq 0$

(2.4b): $1 - p_1 - p_2 < 0$
Equation \( \Delta w_a^2 - (1 - p_1 - p_2 - \varepsilon) \Delta w_a + \frac{p_2(1-p_2)\varepsilon}{p_1-p_2} = 0 \), we find that Item (9) will be negative if \( \Delta w_a \in (\frac{(1-p_1-p_2-\varepsilon) - \sqrt{M}}{2}, \frac{(1-p_1-p_2-\varepsilon) + \sqrt{M}}{2}) \).

To summarize, give any \( p_1, p_2 \) that follows \( 0 < p_2 < p_1 < 1 \), the analyses above suggest that:

\[
\begin{align*}
(1) & \text{OR}_t < \text{OR}_{t+1} \text{ if } \varepsilon \in \left( \frac{(\sqrt{p_1(1-p_1)} - \sqrt{p_2(1-p_2)})^2}{p_1-p_2}, 1-p_1 \right), \text{ or if } \varepsilon \in \left( p_2 - p_1, \frac{(\sqrt{p_1(1-p_1)} - \sqrt{p_2(1-p_2)})^2}{p_1-p_2} \right), \\
& \text{and } \Delta w_a \in \left( 0, \frac{1}{2} \right) \text{ or } \Delta w_a \in \left( \frac{(1-p_1-p_2-\varepsilon) - \sqrt{M}}{2} \right), 1-p_2; \\
(2) & \text{OR}_t = \text{OR}_{t+1} \text{ if } \varepsilon \in \left( p_2 - p_1, \frac{(\sqrt{p_1(1-p_1)} - \sqrt{p_2(1-p_2)})^2}{p_1-p_2} \right), \Delta w_a \in \left( \frac{(1-p_1-p_2-\varepsilon) - \sqrt{M}}{2} \right), 1-p_2; \\
(3) & \text{OR}_t > \text{OR}_{t+1} \text{ if } \varepsilon \in \left( p_2 - p_1, \frac{(\sqrt{p_1(1-p_1)} - \sqrt{p_2(1-p_2)})^2}{p_1-p_2} \right), \Delta w_a \in \left( \frac{(1-p_1-p_2-\varepsilon) - \sqrt{M}}{2} \right), 1-p_2.
\end{align*}
\]

where \( p_2 - p_1 < \varepsilon < 1 - p_1 \), \( 0 \leq \Delta w_a < 1 - p_2 \), \( M = (1 - p_1 - p_2 - \varepsilon)^2 - 4 \cdot \frac{p_2(1-p_2)\varepsilon}{p_1-p_2} \).

Particularly, for the special case of \( \varepsilon = 0 \), \( \text{OR}_{t+1} = \text{OR}_t \) if \( \Delta w_a = 1 - p_1 - p_2 \); \( \text{OR}_{t+1} > \text{OR}_t \) if \( \Delta w_a > 1 - p_1 - p_2 \); \( \text{OR}_{t+1} > \text{OR}_t \) if \( \Delta w_a < 1 - p_1 - p_2 \). In other words, if the new opportunity is evenly distributed between the service and the working classes (\( \Delta s_a = \Delta w_a \)), any \( p_1 < 0.5 \) (\( 1 - p_1 - p_2 > 0 \)) will predict a higher odds-ratio at time \( t+1 \); any \( p_2 > 0.5 \) (\( 1 - p_1 - p_2 < 0 \)) will predict a lower odds-ratio at time \( t+1 \). Substantially, it means that the start values of \( p_1 \) and \( p_2 \) do affect the direction of the change in the odds-ratios.

Figure 2.4 provides an illustrative example. When \( 1 - p_1 - p_2 \geq 0 \), Figure 2.4a shows that any \( \varepsilon \in \left( \frac{(\sqrt{p_1(1-p_1)} - \sqrt{p_2(1-p_2)})^2}{p_1-p_2}, 1-p_1 \right) \) will predict a higher odds-ratio (in maroon) at time \( t+1 \). If \( \varepsilon \in \left( p_2 - p_1, \frac{(\sqrt{p_1(1-p_1)} - \sqrt{p_2(1-p_2)})^2}{p_1-p_2} \right) \), the curve (in black) representing

\[ \Delta w_a = \frac{(1-p_1-p_2-\varepsilon) - \sqrt{M}}{2} \text{ and } \Delta w_a = \frac{(1-p_1-p_2-\varepsilon) + \sqrt{M}}{2} \] indicates a constant odds-ratio before and after the expansion, and the areas in red and in gray give the respective sets of \( \Delta w_a \) and \( \varepsilon \) that predict a higher and lower odds-ratio at time \( t+1 \). When \( 1 - p_1 - p_2 < 0 \), however, Figure 2.4b indicates that any \( \varepsilon \) greater than 0 predict \( \text{OR}_{t+1} > \text{OR}_t \). This is also evident by analyzing Item (9) in which
each part is now greater than 0 (for example, \(-1 - p_1 - p_2 - \varepsilon) \Delta w_a > 0\). In words, if the distribution of the new opportunity disproportionately benefit the service class (\(\Delta s_a > \Delta w_a\)), any start value of \(p_1\) and \(p_2\) with \(1 - p_1 - p_2 < 0\) will predict increasing educational inequality at time \(t+1\).

Together, the conditions in (10) suggest that the direction of the change for the odds-ratio depends on the initial proportions of the S* and W* students in level k schools before the expansion (as captured by parameters \(p_1, p_2\)) and how the new opportunities are distributed between the two groups (as captured by \(\varepsilon\) and \(\Delta w_a\)). If the values of \(p_1\) and \(p_2\) are fixed, then how the odds-ratio changes depends on the relative sizes of \(\Delta s_a\) and \(\Delta w_a\) (as captured by the different sets of \(\Delta w_a\) and \(\varepsilon\) in Figure 2.3).

Net of the resource factor, the relative sizes of \(\Delta s_a\) and \(\Delta w_a\) are further determined by how educational expansion changes the ability requirement of schools from time \(t\) to \(t+1\). Because of the assumed differentials in the ability distribution between the S* and W* groups, when the two groups are combined, the overall population ability turns out to have a bi-modal normal distribution with the mean equals \((\bar{a}_s + \bar{a}_w)/2\), as illustrated in Figure 2.4. Since \(\bar{a}_s > \bar{a}_w\), \(a^0_{(t+1)} < a^0_t\), it follows that:

\[
\begin{align*}
\text{If } (\bar{a}_s + \bar{a}_w)/2 &\leq a^0_{(t+1)} < a^0_t, \Delta s_a > \Delta w_a, \\
\text{If } a^0_{(t+1)} < a^0_t &\leq (\bar{a}_s + \bar{a}_w)/2, \Delta s_a < \Delta w_a, \\
\text{If } a^0_{(t+1)} &< (\bar{a}_s + \bar{a}_w)/2 < a^0_t, \Delta s_a < \Delta w_a \text{ if } (\bar{a}_s + \bar{a}_w)/2 - a^0_{(t+1)} < a^0_t - (\bar{a}_s + \bar{a}_w)/2; \\
\Delta s_a &> \Delta w_a \text{ if } (\bar{a}_s + \bar{a}_w)/2 - a^0_{(t+1)} > a^0_t - (\bar{a}_s + \bar{a}_w)/2 \\
\Delta s_a &> \Delta w_a \text{ if } \frac{\bar{a}_s + \bar{a}_w}{2} - a^0_{(t+1)} = a^0_t - (\bar{a}_s + \bar{a}_w)/2.
\end{align*}
\]
Substantively, it means that the distribution of the newly created schooling opportunity for the S* and W* groups ($\Delta s_a$ and $\Delta w_a$) is contingent on the initial group difference in ability ($\bar{s}_s, \bar{s}_w$), the initial ability requirement ($a^0$, which determines $p_1$ and $p_2$) and the extent of educational expansion (how the ability requirement shifts in the overall ability distribution). Figure 2.5 gives an illustration of the process.

Figure 2.6 indicates the dynamics.

For example, other factors being equal, if schools are only open to the top 25% of the overall population before the expansion ($a^0$), given the initial class differentials in the ability

---

17 One approximation of the ability requirement percentile is the transition rate, which is the conditional probability of students continuing to level k given their completion of level k-1.
Figure 2.6: An Illustrative Distribution of Newly-Created Opportunity between Groups
distribution, the service class students will disproportionally benefit from educational expansion until the schools are open to the top 75% of the overall population. In contrast, if a certain level of education is already prevailing in the overall population, say, the top 75% of the overall population have already been in school, further educational expansion will disproportionally benefit the working class students ($\Delta s_a < \Delta w_a$ in the last panel). As a result, controlling for other factors, if the ability distribution between the S* and W* groups is fixed ($\bar{a}_s, \bar{a}_w$), and if the initial $p_1, p_2$ parameters are also fixed, then each level of the ability requirement determines a respective set of $\Delta w_a$ and $\varepsilon$, which uniquely decides how $p_1, p_2$ change under educational expansion.

Figure 2.6b depicts an example of the changing distribution of $p_1, p_2$ at various levels of ability requirement of schools in a given setting ($\theta=0.5, \bar{a}_w = 1, \delta_{as} = \delta_{aw} = 1$). It shows the paths of $p_1, p_2$ when the expansion starts from the ability requirement as the top 1% (schools only admit the top 1% of the students) to that as the top 99%. Different start values of $p_1, p_2$ therefore return different odds-ratios at time $t$, and different extents of educational expansion produce corresponding odds-ratios in time $t+1$. In this example, if the expansion starts from the top 20% ability requirement (i.e. only admitting students in the top 20%), the odds-ratio will be declining until the schools open to 80% of the overall population of the students. As a result, whether educational inequality is increasing, decreasing or persistent essentially depends on the starting position of the expansion and the time points of our observation.

In the above analyses, I have set $\varphi=1$, meaning that there is no group difference in the distribution of resources. I have demonstrated how the new educational opportunities, produced by the structural change of educational expansion, are distributed across different social groups,

---

18 This is because of the assumption that ability follows a normal distribution (a bell curve).
and how different scenarios of allocations further introduce changes in the odds-ratio. Now, going back to the definition of the probabilities of continuing to level $k$ education for $S^*$ and $W^*$ students in Equation (12), I set $\theta=1$ so that there is no group difference in ability between the service class and the working class students. Similarly, let $p_1' = pr(r_{is} > r^0_l)$, $p_2' = pr(r_{iw} > r^0_l)$, and let $\Delta s_r$ and $\Delta w_r$ be the respective proportions of students in the service and working class that are able to take the new opportunities in the level $k$ education. Similar to (13) and (14), write

$$\text{OR}_t' = \frac{p_1' - p_1}{p_2' - p_1}, \quad \text{and} \quad \text{OR}_{t+1}' = \frac{p_1' - p_1 + (p_1' \Delta s_r + p_2' \Delta w_r + \Delta s_r \Delta w_r) + \Delta s_r}{p_2' - p_1' + (p_1' \Delta s_r + p_2' \Delta w_r + \Delta s_r \Delta w_r) + \Delta w_r}.$$

When the resource requirement is declining under educational expansion (i.e. $r^0_l > r^0_{l+1} \geq 0$), the difference in $\text{OR}_{t+1}'$ and $\text{OR}_t'$ introduced by the resource requirement factor follows exactly the same logic as that of the ability requirement factor. Thus the pattern is completely analogous to the one established based on the ability requirement factor. If the resource requirement is increasing (namely $0 \leq r^0_l < r^0_{l+1}$), however, the proportion of individuals with higher resources than $c$ (minimal resource requirement) gets smaller, net of other factors. That is, $pr(r_i > r^0_l) < pr(r_i > r^0_{l+1})$ if $0 \leq r^0_l < r^0_{l+1}$. In this scenario, the proportion of students from both the $S^*$ and $W^*$ groups who choose to continue education at $k$ level will be smaller at time $t+1$, i.e. $\Delta s_r < 0, \Delta w_r < 0$. The conditions that predict the directions of the odds-ratio change are thereby the opposite of those conditions for when the resource requirement is decreasing (Mathematically, this is proven by setting $\Delta w_r = -\Delta w_r \ln (9)$).

To sum up, analyses in this section show that both structural and institutional changes introduced by educational expansion impact educational inequalities at the $k^{th}$ level. The structural change, in the form of changing ability requirement, determines the extent of availability of new educational opportunities; the institutional change in the form of the changing resource requirement decides who can finally benefit from the newly created educational
opportunities. In the context of such structural and institutional changes, students use their own academic ability and resources to make educational decisions following a relative risk aversion strategy. As a result, class differentials in the distributions of academic ability and resources determine how different social classes are finally matched to the new educational positions in the schools.

**Educational Transition to Level k+1**

My discussion above shows how educational expansion affects patterns of educational inequality at level k. Since educational transition is a multi-stage process, those who have continued to level k education make up the pool of total individuals who can potentially transition into the k+1\textsuperscript{th} educational level (Klugman 2012).

Let level k be the first post-compulsory educational level of interest and level k+1 the next so that only students having completed the k\textsuperscript{th} level education can be admitted to the k+1\textsuperscript{th} level. Suppose that educational expansion takes place from time t to t+1 on both levels with the rates of E\textsubscript{k} and E\textsubscript{k+1}, respectively. A larger rate of educational expansion indicates a faster change in the ability requirement percentile of schools for the overall population.

To more easily show the linkage of the inequality patterns at the two levels, I consider the two following simple scenarios. First, holding the resource requirement constant across groups and across educational levels for now, set E\textsubscript{k+1} as 0 and E\textsubscript{k} as greater than 0. In words, this means that educational expansion happens at level k but not at level k+1, i.e. the level of ability requirement is declining at level k but is held constant at level k+1. Let \( e_j = \Pr(a_{ij} > a_{k+1,j,t}) \), \( f_j = \Pr(a_{k,j,t} < a_{ij} < a_{(k+1),jt}) \) for \( j = S^* \) and \( W^* \) groups, which respectively indicate the proportions of level k students who exceed the ability requirement for level k+1 and those who do not at time t. Let \( \Delta f_j \) denote the increased proportion of students who go to level k for the \( j^\text{th} \) group at time t+1.
Thus for the S* group, \( p_{1,(k+1),t} = e_s/(e_s + f_s) \), \( p_{1,(k+1),(t+1)} = e_s/(e_s + f_s + \Delta f_s) \); for the W* group, \( p_{2,(k+1),t} = e_w/(e_w + f_w) \) and \( p_{2,(k+1),(t+1)} = e_w/(e_w + f_w + \Delta f_w) \). The difference in odds ratio at level k+1 between time \( t \) and \( t+1 \) is then defined as:

\[
\text{OR}_{k+1,(t+1)} - \text{OR}_{k+1,t} = \frac{e_s}{e_s+f_s} \left( 1 - \frac{e_s}{e_s+f_s} \right) \Delta f_s,
\]

Since \( \Delta f_s > 0 \) and \( \Delta f_w > 0 \), \( \text{OR}_{(k+1),(t+1)} \) will always be greater than \( \text{OR}_{(k+1),t} \). This suggests that if educational expansion takes place for level k education but not for level k+1 education, the odds-ratio will always be increased at level k+1.

Secondly, if \( E_k \) is 0 whereas \( E_{k+1} \) is greater than 0, educational expansion takes place at level k+1 but not at level k. In other words, there is a lower ability requirement for schools at level k+1. The proportions for the S* and W* students entering level k+1 education are now defined as: \( p_{1,(k+1),t} = e_s \), \( p_{1,(k+1),(t+1)} = e_s + \Delta f_s \), \( p_{2,(k+1),t} = e_w \), \( p_{2,(k+1),(t+1)} = e_w + \Delta f_w \). The calculation of odds-ratios is then the same as that in (8), but with the constraint that a student’s ability must exceed the minimum ability requirement for the level k education. I will not go to the details of how the relative sizes of \( \Delta f_s \) and \( \Delta f_w \) and the initial proportions \( e_f \) and \( e_w \) determine the change in the odds-ratios, because the conditions are similar to those in (10).\(^{19}\)

A more general version of the two scenarios above is when neither \( E_k \) nor \( E_{k+1} \) is 0, yet they are both positive from time \( t \) to \( t+1 \). This means that educational expansion happens on both the k\(^{th}\) and the k+1\(^{th}\) levels, but with varying rates. Since the rate of educational expansion is a

\(^{19}\) Detailed proof is available upon request.
direct indicator of the change in schools’ ability requirement percentiles for the overall population, the relative sizes of the rates at level k and level k+1 will indicate the different gaps of the changing ability requirement percentiles (e.g., from the top 25% to the top 50%) at both levels. Specifically, net of other factors, if $E_k > E_{k+1}$, this scenario can be reduced to the first scenario discussed above and we will always observe increased odds-ratio at level k+1. Similarly, if $E_k < E_{k+1}$, it can be reduced to the second scenario discussed above. Here, the changes in odds-ratio at level k+1 will be determined by the sign of Item (12), conditioning on $a_i \in \min(a_{k,t}, a_{k,(t+1)})$, $\min(0, a_{k,t}, a_{k,(t+1)})$, $+\infty$. Finally, if $E_k = E_{k+1}$, educational expansion happens in the same rate on both level k and level k+1, and the increased number of people entering level k education will be completely absorbed by the increased slots in level k+1, Therefore the odds-ratios at both levels will stay unchanged. As a result, the rates of educational expansion at both levels have impacts on how the origin-education association at level k+1 is associated with that at level k.

Similar to previous analyses, since analytically the ability requirement factor and the resource requirement factor are interchangeable if the resource requirement is declining, the patterns illustrated above also hold if we examine the effects of the resource requirement factor while holding the ability distribution constant between groups S* and W* (i.e. $\theta = 1$). If the resource requirement is increasing, again, the direction of the change in the odds-ratio will be the opposite of the patterns illustrated above.

Moreover, as defined in (5), the $\theta$ and $\varphi$ parameters capture the differences in the class differentials with respect to academic ability and resources among different societies. Breen et al (2009) speculate that if the class differentials in academic ability become smaller, it tends to shrink educational inequalities. Analogously, since $0 < \theta, \varphi \leq 1$, it is plausible that net of others, a
smaller value of the $\theta$ and $\phi$ parameters will inflate educational inequalities at level $k$ and level $k+1$.

**Illustrative Simulation**

To empirically demonstrate how various extents of educational expansion produce different patterns of educational inequality, I conduct the following simulation analyses.

First, I draw two random samples with $N=100,000$ each, representing the two groups, $S^*$ and $W^*$. I then create three random variables for each group: ability ($a$), subjective probability of success ($\pi$), and resource ($r$). Both the ability parameter and the subjective probability of success parameter follow a normal distribution with $\delta = 1$ for each group, but with different group means ($\mu_{aw} = \theta^* \mu_{sw}$). Given that the two variables $a$ and $\pi$ are highly correlated, I assign a correlation coefficient 0.85 to them.\(^\text{20}\) The resource variable follows a logistic distribution\(^\text{21}\) with same dispersion for each group, while the means are different, i.e. $\mu_{rw} = \phi^* \mu_{rs}$. The parameters $\theta$ and $\phi$ capture the group differences in the means of the ability and resource distributions.

Next, I pool the two groups together, and set various selection criteria at both the $k^{\text{th}}$ level and the $k+1^{\text{th}}$ level. I use the percentiles of the pooled sample ($N=200,000$) characteristics (specifically, $a$ and $r$, since $\pi$ is highly correlated with $a$) to simulate different extents of educational expansion. For instance, if the two groups have identical resource distribution ($\phi = 1$) but different means in ability (say, $\theta = 0.8$, $\mu_{aw} = 0.8^* \mu_{sw}$), changing the selection criteria at both

\(^{20}\) The choice of the correlation coefficient is arbitrary, but the empirical patterns are similar when using different values of the correlation coefficient.

\(^{21}\) In practice, to easily compare the means across groups, I draw a log-normal distribution (normal-like) instead of the logistic distribution for this random variable.
Table 2.2: Simulation Results of Educational Expansion (N=200,000)

<table>
<thead>
<tr>
<th>Educational expansion at level k</th>
<th>Odds-ratio at level k</th>
<th>Educational expansion at level k+1, odds-ratio at level k+1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>75th percentile</td>
<td>50th percentile</td>
</tr>
<tr>
<td>(\theta=0.8, \varphi=1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>60th percentile</td>
<td>1.35</td>
<td>1.24</td>
</tr>
<tr>
<td>40th percentile</td>
<td>1.29</td>
<td>1.29</td>
</tr>
<tr>
<td>(\theta=0.5, \varphi=1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>60th percentile</td>
<td>2.14</td>
<td>1.75</td>
</tr>
<tr>
<td>40th percentile</td>
<td>1.90</td>
<td>1.91</td>
</tr>
<tr>
<td>(\theta=1, \varphi=0.8)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>60th percentile</td>
<td>1.30</td>
<td>1.21</td>
</tr>
<tr>
<td>40th percentile</td>
<td>1.24</td>
<td>1.25</td>
</tr>
<tr>
<td>(\theta=1, \varphi=0.5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>60th percentile</td>
<td>1.95</td>
<td>1.60</td>
</tr>
<tr>
<td>40th percentile</td>
<td>1.72</td>
<td>1.76</td>
</tr>
</tbody>
</table>

Note: The larger the percentile, the more selective the school admission is. Odds-ratio is calculated by \(\frac{p_1/(1-p_1)}{p_2/(1-p_2)}\).

The kth level and the k+1th level from the top 1%, 25%, 40%, 50%, 60%, 75%, to the 99% percentile of the ability variable produces different extents of educational expansion.

The simulated results of the odds-ratios are reported in the first panel of Table 2.2. To save space, it only reports the odds-ratios based on the selection criteria as the 40% and 60% percentiles at level k and those as the three quartiles at level k+1. Note that the larger the percentile, the more selective the school admission is.\(^{22}\)

The first panel shows that as the level k education (say, senior high school) incorporates more individuals, e.g. as the selection criteria on this level move from admitting the top 40% to admitting the top 60%, the odds-ratio at level k is declining from 1.35 to 1.29. In the meanwhile, as expected, the odds-ratio at level k+1 (say, college) is increasing if the ability requirement is fixed at level k+1 but becomes lower at level k. For example, in the case that the level k+1 education opens to the top 25% of the students who complete the level k education, the odds-

\(^{22}\) The 60% percentile selection criterion means that schools only choose those whose ability is above the 60% percentile of the overall ability distribution. In other words, education is only open to the top 40% of all the students in the two groups.
ratio on the k+1 level increases from 1.24 to 1.29 as the level k education is available from the top 40% to the top 60% of the overall student population.

Comparing the first and the second panels with equal $\varphi$ and different $\theta$’s, we again find that as expected, the smaller $\theta$ is, the larger the odds-ratios are. The third and fourth panel shows that the same pattern applies to $\varphi$ when $\theta$ is set to be 1 (equal ability).

Overall, Table 2.2 supports that there is a linkage between the inequality patterns at the level k and the level k+1 education. Specifically, a disproportionally increased supply of students at the level k implies a higher odds-ratio at level k+1. Furthermore, different sets of the $\theta$ and $\varphi$ parameters drive the differences of educational inequality patterns among different societies.

**Educational Expansion and Education-related Labor Market Outcomes**

Educational expansion affects not only the process of educational attainment, but also the quantity and composition of skilled workers in the labor market. Various extents of educational expansion channel different sets of individuals into schools, and thus supply different quantities of educated labor with varying average qualities.

**Origin-Education-Destination Associations**

Educational expansion has impacts on the origin-education-destination relationships in two ways: firstly, it directly affects the origin-education relationship; secondly, it influences how education mediates the linkage between origin and destination because of the increased supply of educated workers. Previous research generally suggests that the association between education and occupation is dependent on the institutional arrangement of particular educational systems and how education is viewed in the labor market (e.g., Shavit and Muller 1998). In other words,
the education-destination relationship is contingent on the institutional settings of a particular society’s educational system and labor market regulations.

If the relationship between education and destination is held as constant, however, the origin-destination linkage is then determined by the origin-education association (Breen 2010a). In this case, Breen and Jonsson (2007) distinguish two effects: first, declining educational inequality indicates a weakening association between social origin and destination, and vice versa (an “equalization effect”). Secondly, if educational expansion drives an increasing share of individuals into a certain level of education and educational inequality at that particular level is also increasing, the average association between origins and destinations will enlarge (a “compositional effect”). However, when and how the two effects hold in the context of educational expansion is not completely clear. Particularly, given that the expansion may occur at various levels of education with different rates, how and why there exit the two effects at each of the different educational levels are yet to be fully theorized.

Formally, assume that the educational system is consisted of two educational levels, i.e. level k and k+1. Let \( Z_k \) and \( Z_{k+1} \) denote the number of students in each level at time t. Educational expansion drives more individuals into each level with different rates, \( E_k \) and \( E_{k+1} \). Therefore, at time \( t+1 \), the increased number of students at both levels is \( \Delta Z_k = Z_k \cdot E_k, \Delta Z_{k+1} = Z_{k+1} \cdot E_{k+1}, \) respectively. Let \( q_1 \) and \( q_2 \) respectively denote the origin-education associations for individuals with level k and level k+1 education at time t; at time \( t+1 \), the origin-education associations become \( q_1' \) and \( q_2' \) where \( \Delta q_k = q_1' - q_1 \), and \( \Delta q_{k+1} = q_2' - q_2 \). \( \Delta q_k \) and \( \Delta q_{k+1} \) measure the changes in the extent of the origin-education association at each level, and a parameter \( v \) captures the difference between them: \( \Delta q_{k+1}/\Delta q_k = v \) when \( \Delta q_k \neq 0, \Delta q_{k+1} \neq 0. \)
Let $\Phi_t$ and $\Phi_{t+1}$ be the overall origin-education associations at time $t$ and $t+1$ respectively. Here, $\Phi_t = \frac{Z_k}{Z_k+Z_{k+1}}*q_1 + \frac{Z_{k+1}}{Z_k+Z_{k+1}}*q_2$, $\Phi_{t+1} = \frac{Z_k+\Delta Z_k}{Z_k+Z_{k+1}+\Delta Z_k+\Delta Z_{k+1}}$.

$q_1' + \frac{Z_{k+1}+\Delta Z_{k+1}}{Z_k+Z_{k+1}+\Delta Z_k+\Delta Z_{k+1}}*q_2'$. Then $\Phi_{t+1} - \Phi_t = \frac{(Z_k+\Delta Z_k)*q_1' + (Z_{k+1}+\Delta Z_{k+1})*q_2'}{Z_k+Z_{k+1}+\Delta Z_k+\Delta Z_{k+1}} - \frac{Z_k*q_1+Z_{k+1}*q_2}{Z_k+Z_{k+1}}$

$= \frac{\Delta q_k*(Z_k+\Delta Z_k+uZ_{k+1}+v\Delta Z_{k+1})}{Z_k+Z_{k+1}+\Delta Z_k+\Delta Z_{k+1}}$ if $\Delta q_k \neq 0, \Delta q_{k+1} \neq 0$. The sign of $\Phi_{t+1} - \Phi_t$ thereby depends on the sign of the following item:

$$\Delta q_k * (Z_k + \Delta Z_k + uZ_{k+1} + v\Delta Z_{k+1}). \quad (13)$$

If $\Delta q_k > 0, \Delta q_{k+1} > 0$, then $v > 0$, the sign of Item (19) is positive, meaning that $\Phi_{t+1} > \Phi_t$. Similarly, if $\Delta q_k < 0, \Delta q_{k+1} < 0$, then $v < 0$, the sign of Item (19) is negative, meaning that $\Phi_{t+1} < \Phi_t$. Furthermore, if $\Delta q_k < 0, \Delta q_{k+1} > 0$, then $v < 0, \Phi_{t+1} > \Phi_t$ if $u < -\frac{Z_k*(1+E_k)}{Z_{k+1}*(1+E_{k+1})}$; $\Phi_{t+1} < \Phi_t$ if $v = -\frac{Z_k*(1+E_k)}{Z_{k+1}*(1+E_{k+1})}$; $\Phi_{t+1} < \Phi_t$ if $u > -\frac{Z_k*(1+E_k)}{Z_{k+1}*(1+E_{k+1})}$.

Finally, if $\Delta q_k > 0, \Delta q_{k+1} < 0$, then $v < 0, \Phi_{t+1} > \Phi_t$ if $u > -\frac{Z_k*(1+E_k)}{Z_{k+1}*(1+E_{k+1})}$; $\Phi_{t+1} < \Phi_t$ if $v < -\frac{Z_k*(1+E_k)}{Z_{k+1}*(1+E_{k+1})}$. In addition, if $\Delta q_k = 0, \Phi_{t+1} - \Phi_t = \frac{\Delta q_{k+1}*(Z_{k+1}+\Delta Z_{k+1})}{Z_k+Z_{k+1}+\Delta Z_k+\Delta Z_{k+1}}, \Phi_{t+1} < \Phi_t$ if $\Delta q_{k+1} < 0; \Phi_{t+1} > \Phi_t$ if $\Delta q_{k+1} > 0$.

Substantively, with a fixed education-destination association, the formulations above suggest that: first, the equalization effect is only possible if educational inequality is persistent at one level but declining at the other levels. Secondly, there is no compositional effect if the odds between the extent of increasing educational inequality at level $k$ and the extent of decreasing inequality at level $k+1$ equal a particular value $v$ ($u = -\frac{Z_k*(1+E_k)}{Z_{k+1}*(1+E_{k+1})}$). Thirdly, a compositional effect is most evident when educational inequalities at various levels change in the same direction (in terms of increasing or decreasing). For example, if at both levels educational
inequality is increasing, then the compositional effect will shift the origin-destination association up.

Finally, if the directions of the change in educational inequality at each level are opposite from each other, then how the compositional effect introduces higher or lower origin-education associations depends on the odds of the change in origin-education associations at various levels ($\mathbf{v}$), the initial share of students at different levels ($\frac{z_k}{z_{k+1}}$), and the rates of educational expansion at each level ($E_k$ vs. $E_{k+1}$). In this regard, educational expansion does have an indirect effect on the origin-destination association.

*Income Returns to Education*

Educational expansion changes the composition of skilled labor in the labor market. On one hand, with the expansion, an increasing number of individuals with higher educational credentials enter the labor market. If the demand for skilled workers does not change, an increasing supply of workers with higher educational credentials tends to lower the price of education (Katz and Murphy 1992).

On the other hand, since the ability requirement of schools lowers under educational expansion, the average cohort quality declines as compared to the cohorts before the expansion. If schools do not compensate the declining cohort quality by providing better education, educational credential as a signal device (Spence 1973; Stiglitz 1975) will then indicate less ability and productivity for cohorts after the expansion, also leading to lower educational returns.

Because of the two effects outlined above (the pricing effect and the signaling effect), educational expansion negatively impacts the income returns to education for cohorts who enter the labor market after the expansion. However, such negative effects hold only as aggregate-level effects. Since income returns to education are also shaped by individual-level factors such...
as ability and resources, the general negative impact educational expansion brings is compensated differently for various social groups.

Again assume that there are two social classes from which students originate: S* denotes the service class and W* denotes the working class. Following the previous notation, let \( p_s \) be the proportion of students completing level \( k+1 \) (say, college) education and let \( 1-p_s \) be the proportion of students with a terminal level \( k \) (say, high school) education at time \( t \), and similarly for \( p_w \) and \( 1-p_w \). Furthermore, denote \( \Delta s \) and \( \Delta w \) respectively as the new proportions of the S* and W* students who enter educational level \( k+1 \) at time \( t+1 \). Therefore, the proportion becomes \( (p_s + \Delta s) \) at level \( k+1 \) and \( (1-p_s - \Delta s) \) at level \( k \) for S* students, and \( (p_w + \Delta w) \) at level \( k+1 \) and \( (1-p_w - \Delta w) \) at level \( k \) for W* students.

Since return to education is shaped also by individual level factors, different classes may have different rates of income returns to level \( k+1 \) education. Here, let \( u_s \) and \( u_w \) indicate the income premium (defined as the mean income gap between students with level \( k+1 \) and level \( k \) education; Goldin and Katz 1999) for the S* and W* students respectively at time \( t \). Similarly \( u_s' \) and \( u_w' \) denote the income premiums for students of the two classes at time \( t+1 \). Let \( \Delta u_s = u_s' - u_s \), and \( \Delta u_w = u_w' - u_w \), which measure the changes in the income premiums for the respective class. Let a parameter \( \eta \) capture the difference between \( \Delta u_s \) and \( \Delta u_w \), namely, \( \Delta u_w = \eta \times \Delta u_s \) \((\Delta u_s \neq 0, \Delta u_w \neq 0)\).

Using similar techniques as from the previous section, I define the mean income premium at time \( t \) to be \( I_t = \frac{p_s}{p_s + p_w} \times u_s + \frac{p_w}{p_s + p_w} \times u_w \), and the mean income premium at time \( t+1 \) to be \( I_{t+1} = \frac{p_s + \Delta s}{p_s + p_w + \Delta s + \Delta w} \times u_s' + \frac{p_w + \Delta w}{p_s + p_w + \Delta s + \Delta w} \times u_w' \). Therefore \( I_{t+1} - I_t = \frac{\Delta u_s \times (p_s + \Delta s + \eta p_w + \eta \Delta w)}{p_s + p_w + \Delta s + \Delta w} \) if \( \Delta u_s \neq 0 \), \( \Delta u_w \neq 0 \). As previously established, returns to education decline in the context of educational expansion, thus \( I_{t+1} - I_t < 0 \), which further produces:
\[ \Delta u_s \times (p_s + \Delta s + \eta p_w + \eta \Delta w) < 0. \] (14)

Note that if \( \Delta u_s > 0 \) and \( \Delta u_w > 0 \), which also implies that \( \eta > 0 \), the inequality in (20) does not hold. If \( \Delta u_s > 0 \) and \( \Delta u_w < 0 \), the income premium becomes larger for the service class but smaller for the working class. Inequality (20) holds only if \( \eta < \frac{p_s + \Delta s}{p_w + \Delta w} \). Since \( p_s > p_w \), \( \eta \) will always be less than -1 if \( \Delta s \geq \Delta w \). In this case, \( \Delta u_s \frac{1}{\eta} \times \Delta u_w < -\Delta u_w \). In words, this means that if the service class has a higher percentage of students who are able to obtain the newly created schooling opportunities than the working class \( i.e. \Delta s \geq \Delta w \), the amount of the enlarged income premium for the service class is always smaller than the amount of the decreased income premium for the working class.

Similarly, if \( \Delta u_s < 0 \) and \( \Delta u_w > 0 \) \( (\eta < 0) \), the income premium is decreasing for the service class but increasing for the working class. Inequality (20) holds only if \( \eta > -\frac{p_s + \Delta s}{p_w + \Delta w} \). \( \eta \) will always be greater than -1 if \( p_s + \Delta s < p_w + \Delta w \), i.e. \( 0 < p_s - p_w < \Delta w - \Delta s \). In this situation, \( \Delta u_w \frac{1}{\eta} \times \Delta u_s < -\Delta u_s \), meaning that the amount of the reduced income premium for the service class is larger than the amount of the increased income premium for the working class. However, if \( p_s + \Delta s < p_w + \Delta w \), \( 1-(p_s + \Delta s) > 1 - (p_w + \Delta w) \), which suggest that educational expansion has greatly benefitted the working class so that, as compared to the percentage of service class students, there even exists a higher percentage of working class students at level \( k+1 \) (or a lower percentage of working class student remaining at educational level \( k \)). This does not seem to be a probable condition. One real world example may be the situation during China’s Maoist era where there exited an egalitarian educational policy that explicitly aimed for benefiting worker and peasant families (Unger 1982).
Further, if $\Delta u_s < 0$ and $\Delta u_w < 0$ ($\eta > 0$), Inequality (14) always holds. Let $\Delta l = l_{t+1} - l_t < 0$, we get: $\eta = 1 + \left( \frac{\Delta l}{\Delta u_s} - 1 \right) \left( 1 + \frac{p_s + \Delta s}{p_w + \Delta w} \right)$. Therefore, $\eta > 1$ if and only if $\Delta l < \Delta u_s < 0$, indicating a smaller decline in educational returns for the service class than for the overall group.

In other words, returns to education decline more for the working class than for the service class. Moreover, as the size of $\eta$ is also determined by $\frac{p_s + \Delta s}{p_w + \Delta w}$, a larger $\frac{p_s + \Delta s}{p_w + \Delta w}$ indicates a larger $\eta$, which further points to greater decline in returns for the working class ($\Delta u_w = \eta * \Delta u_s < 0$). In this sense, different sets of $p_s, p_w, \Delta s$ and $\Delta w$ can produce different values of $\eta$, pointing to various levels of groups differentials in the declining returns to education.

As illustrated in (11), educational expansion determines how the increasing educational opportunity is distributed across groups, i.e. the relative sizes of $\Delta s$ and $\Delta w$. Therefore, educational expansion indeed produces a compositional effect for income returns to education. With varying extent of educational expansion, the compositional effect can be different.

**Summary**

To summarize, in this section, I have studied three sequential processes of how educational expansion affects education-based inequality and mobility: educational attainment at level k, educational attainment at level k+1, and how the volume and composition of educated workers have impacts on the origin-education-destination associations and on income returns to education. Through formal and simulation analyses I have shown that these processes are intrinsically linked to one another.

Specifically, my analyses document three main results. First, I find that how educational expansion affects educational attainment depends on three factors: (1) initial group differentials in the ability and resource distribution ($\theta$ and $\varphi$ parameters); (2) each group’s proportion of
continuing education before the expansion ($p_1$ and $p_2$); and (3) how new educational opportunities are distributed among different groups ($\Delta s$ and $\Delta w$). The $\theta$ and $\varphi$ parameters capture the group differentials in terms of ability and resource within a society. With different $\theta$’s and $\varphi$’s, patterns of educational inequalities in different societies vary. The $p_1$ and $p_2$ parameters capture the initial group shares in schools, which also reflect the ability and resource requirements of schools before educational expansion. The $\Delta s$ and $\Delta w$ parameters represent how different social groups are able to benefit from educational expansion. They are shaped by the structural and institutional changes introduced by educational expansion. As Figure 2.3 shows, different sets of $\Delta s$ and $\Delta w$ lead to various patterns of educational inequalities.

Secondly, I find that the relative sizes of educational expansion rates at different levels ($E_k$ and $E_{k+1}$) affect the linkage between the inequality patterns at the $k^{th}$ and $k+1^{th}$ levels. My simulation analyses show that educational inequality in level $k+1$ is associated with that in level $k$. In this sense, educational expansion generates sequential consequences at different levels.

Thirdly, educational expansion indirectly affects education-related labor market outcomes through a potential compositional effect. The compositional effect for the origin-destination association depends not only on the ratio of students at the level $k$ over $k+1$, but also on the relative sizes of the educational expansion rates at the two levels ($\eta$). The compositional effect for the changing income premiums of the $k+1^{th}$ level education is contingent upon the volume and composition of educated workers at that level. Specifically, the change in the premiums is dependent on the group composition in schools after the expansion ($\eta$).

Overall, based on these findings, I argue that educational expansion is a dynamic process that takes place on multiple levels of schooling. As the expansion occurs on different educational level with varying rates, it produces level-specific effects regarding education-based inequalities.
that are sequentially linked. Moreover, the exact patterns of educational inequalities are shaped by both individual and structural/institutional level factors. While the structural change determines the general availability of the new schooling opportunities, the institutional change decides to whom the new opportunities are attainable. The structural and institutional changes make up the new budget constraints under educational expansion within which rational actors make conscious decisions based on their available ability and resources. In addition, the results also show that educational expansion impacts not only the processes of educational attainment, but also the inequality patterns in education-based labor market outcomes, specifically, occupational opportunities and income returns. In this regard, educational expansion affects education-based inequality and mobility step-by-step.

Furthermore, these findings also shed light upon why there exist cross-country empirical variations in the education-based inequality and mobility patterns under the expansion. This is because national-specific institutional settings influence the initial class differentials in the ability and resource distributions ($\theta$ and $\varphi$ parameters) as well as how the educational system works. Specifically, how the educational system is set up in each country produces not only variations in the class compositions in schooling before the expansion ($p_1$ and $p_2$), but also variations in the nature of educational expansion per se, i.e. at which educational level the expansion takes place and the corresponding rates, as well as the structural and institutional changes that determine who are finally able to truly benefit from the expansion ($\Delta s$ and $\Delta w$).

Figure 2.7 provides a summary of the above theoretical model and respective parameters.
Figure 2.7: A Summary of Theoretical Modeling and Parameters
It suggests that the element of nation-specific institutional contexts directly determines the initial distributions of ability and resources among social groups ($\theta$ and $\varphi$). These contexts also determine the proportion of students in each social group that have successfully attained a certain level of education ($p_1$ and $p_2$). In addition, these contexts indirectly affect how the structural and institutional changes introduced by educational expansion shape the distribution of the newly created educational opportunities among various groups ($\Delta s$ and $\Delta w$). Together, the six parameters affect individuals’ educational attainment at both the $k^{th}$ and the $k+1^{th}$ levels.

Moreover, there is a linkage between $k^{th}$ and $k+1^{th}$ educational levels. Various relative rates of educational expansion at both levels (as captured by $E_k$ and $E_{k+1}$) produce different configurations of inequality patterns. Finally, the inequality patterns with respect to educational attainment further impact education-related labor market outcomes by a compositional effect, which is captured by the $\nu$ and $\eta$ parameters respectively for the origin-education-destination associations and income returns to education.

**CONCLUSION**

To conclude, in this chapter I formally proposes a general theory of educational expansion to explain how and why educational expansion can produce various patterns of education-based inequality and mobility across country and time. By integrating both micro- and macro-level theoretical approaches, I have examined the relationship between educational expansion and educational attainment, and how educational expansion affects the origin-destination associations as well as income returns to education. I argue that to understand why there exist different empirical patterns in various societies, we need to regard educational
expansion as a dynamic process that produces multi-leveled intrinsically connected consequences.

On one hand, educational expansion has direct impacts on how people are socially selected into schools. On the other hand, the group differentials in the schooling process subsequently transmit into a compositional effect on occupational and material returns to education. In other words, how individuals make educational choices under the expansion is essentially linked to how their choices are later rewarded in the labor market. This connection holds because individuals are conscious decision makers in the process of education-based social selection and social differentiation: individuals make certain educational decisions based on the ability and resources they have and adjust their decisions according to the changing macro-level structural and institutional constraints in the educational system. Given that the distributions of ability and resources are different across social groups, individuals may have differential preferences for educational attainment and subsequently adopt varying strategies for social mobility. Such group variations can further be moderated by educational expansion, as educational expansion modifies the educational opportunity structure and shifts the selection rules of schooling, thereby introducing changes in the ability and the resource requirements of schools. Different extents of educational expansion pose different demands for individuals’ ability and resources, so that the availability of the new schooling opportunities and the way in which the newly-created opportunities are distributed among social groups can vary, which further produce various patterns of education-based inequality and mobility.

In this regard, based on formal modeling and simulation analyses, I make two sets of arguments. First, I argue that educational expansion exerts stepwise effects on education-based inequality and mobility. My analyses have shown that the inequality pattern for a given post-
compulsory level of education is linked to that of the subsequent level. Particularly, depending on the relative rates of educational expansion at two adjacent levels, the linkage of the educational inequality patterns at both levels can be increasing or declining. My formulation also shows that educational expansion can have a compositional effect on the origin-education-destination associations and the income returns to education. In particular, how the new educational opportunities are distributed among social groups affects the composition of skilled labor in the labor market.

Secondly, I argue that how educational expansion affects education-based inequality and mobility is contingent upon the following factors in a given setting: (1) the initial social group differences in the ability and resource distribution; (2) the group share of schooling at a certain level before the expansion; (3) the structural and (4) institutional changes introduced by educational expansion that modify the ability and resource requirements for schooling. Across societies, the institutional differences in the sociopolitical contexts further drive the variations in the patterns caused by the above factors.

In the following chapters, I will use the case of reforming China to empirically demonstrate and test the arguments above. I will show that with its particular structural and institutional changes introduced by educational expansion, China’s patterns of educational inequality, social mobility, and income returns to education are different from those in industrialized societies.
Chapter 3

Educational Expansion and Its Consequences in Reforming China: Background and Hypotheses

INTRODUCTION

While educational expansion takes place in many societies, its consequences in terms of education-based inequality and mobility are not always the same. As my previous chapter suggests (Figure 2.2, Chapter 2), nation-specific institutional contexts (such as the initial conditions of the social structure and the schooling system) and the institutional and structural changes introduced by educational expansion together determine the patterns of educational inequality, which further impact education-related labor market outcomes after the expansion. In this dissertation, to demonstrate the general theory of educational expansion proposed in Chapter 2, I focus on the processes and consequences of educational expansion in reforming China.

Reforming China is one especially interesting case for studying the consequences of educational expansion. Over the past three decades, there have been several waves of educational reforms in the Chinese society (Ministry of Education, 2011). Particularly, in 1999 there was a sudden expansion of schooling opportunities at upper secondary levels (Ministry of Education, 1998), making China a site of “natural experiment” (Dunning 2008; Meyer 1995) for social studies. Yet, within the discipline of sociology, China’s experience of educational expansion in the last thirty years has rarely been included in comparative education-based stratification research (with only a few exceptions such as Wu and Zhang 2010; Zhou 1998), and the specific institutional and structural characteristics in China’s educational expansion process have not been widely recognized and systematically discussed. Therefore, an examination of the case of reforming China under a comparative framework will provide a valuable chance to build up
stylized facts not only to understand the consequences of China’s educational expansion, but also to further develop the theoretical accounts in the fields.

In this chapter, I will give the necessary background information of China’s educational expansion in the reform period. I will first briefly introduce China’s educational system, and then illustrate China’s unique structural and institutional features under educational expansion. Next, I will discuss how educational expansion has affected educational attainment and education-related labor market outcomes. Using the general theory in Chapter 2, I will also hypothesize the consequences of educational expansion based on China’s particular institutional contexts.

EDUCATIONAL EXPANSION IN CHINA’S REFORM PERIODS

China’s Educational System

Like many other societies, China’s full-time education ranges from primary school, lower secondary school (junior middle school), and upper secondary school (academic and vocational senior high school), to tertiary education (junior college and university). Since the release of the 1986 Compulsory Education Law (Ministry of Education, 1986), nine-year education (usually 6 years of primary school plus 3 years of junior middle school) has been designed as compulsory and free. However, for much of the ensuing period, free compulsory education was still a target rather than a realized goal throughout China (Tsui 1997). Schools usually charged a small amount of money for primary and lower secondary education, especially in rural areas before 2007. As many poor families had difficulty paying school fees, some children were forced to quit earlier than the nine-year goal. It was not until 2011 that the goal of universalizing compulsory education had been achieved (Ministry of Education, 02/09/2012).

The post-compulsory education starts from upper secondary school, which encompasses
an academic track (academic senior high school – *gaozhong*) and a vocational track (including technical school, associate school and vocational high school – *zhongzhuang, zhigao*, and *jixiao*).

After the completion of compulsory education, students who wish to continue their education need to take the *High School Entrance Exam (zhongkao)*, which is administered by provincial level Education Examinations Authorities (*jiaoyu kaoshiyuan*, an agency under the Ministry of Education). The exam is standardized within provinces, and determines whether students are eligible for the academic track or the vocational track. The academic track not only requires higher *zhongkao* exam scores for admission, but also directly equips students for tertiary education. Therefore, it is widely considered as a structural gateway for college (Li 2010; Shen 2003; Min 2007; Yang 2006). The vocational track usually emphasizes specific trade skills such as nursing, auto mechanic, and electronic engineering, and is the less academic of the two.

Typically, students will need to complete the academic senior high school to take the *College Entrance Exam (gaokao)*. Vocational upper secondary school graduates are also allowed to take that exam if their academic ability is considered to be “equivalent” to senior high school graduates’, but the proportion of students from vocational school among all the exam takers is small: normally less than 5 percent (*People’s Daily Online*, 02/09/2011).

Admission to higher education is predominantly decided by students’ performance on the *College Entrance Exam*. Before 2002, the exam was monitored by the National Education Examinations Authorities and was nearly nationally standardized (except for *Shanghai*). Since 2004, a few provinces have been authorized to autonomously decide the test subjects and exam materials. Admission to college is operated within each province based on national enrollment

---

23 According to two national representative data sources that provide detailed educational history information (the 1996 *Life History and Social Change Survey* and the 2008 *China General Social Survey*), the percentage of vocational school graduates successfully entering college was 4.85 percent in 1996 and 4.98 percent in 2008.
planning. On a yearly basis, the Ministry of Education determines the total number of college enrollments and distributes the enrollments by provinces for each university (Dong and Xu 2009). When applying for college, students need to submit a list (zhiyuan form) prioritizing their choices of colleges and majors based on their estimated or real scores on the College Entrance Exam. The provincial educational authorities then match students to colleges and majors according to students’ choices and exam scores as well as the provincial enrollment plans (Wang 2012). Each year the Ministry of Education also determines the cut-off points of exam scores (luqu fenshuxian) as the minimum entry requirements for universities (benke) and junior colleges (dazhuan) for each province (Hannum et al 2012). Universities are regarded as more selective and prestigious than junior colleges, and take longer to complete (usually four rather than three years).

Figure 3.1 summarizes the hierarchy and sequence of schooling progression in China’s educational system, with arrows indicating the transitions. It suggests that the high school and college entrance exams play an important role in determining transitions to upper secondary and
tertiary education, as well as the type of education one receives. Through the entrance exams, educational authorities control the yearly national and provincial enrollment for schools based on national enrollment planning. Therefore, if the state has a target enrollment rate for schools, it could realize it by changing the school enrollment plans. In this sense, the national enrollment planning system is one important feature to understand China’s educational expansion.

**China’s Educational Expansion**

Over the past three decades, China’s educational system has undergone significant changes. Along with China’s successful economic development, the Chinese government has dramatically increased educational resources and expanded education at all levels. A great deal of progress has been made with respect to the enforcement of compulsory education. According to official reports, in 2000, over 85 percent of the population has achieved the policy goal of “basically universalized nine years of compulsory education and basically eradicated illiteracy among the youth” (Xinhua Net, 26/11/2007). By the end of 2011, the Chinese government claims that China has reached the goal of universalizing compulsory education (Ministry of Education, 09/10/2012), as noted earlier.

There were also efforts to advance upper secondary level education. In the mid-1990s, the policy aim was to “actively develop vocational education, and moderately enlarge the scale of senior high school education” (Ministry of Education, 1993, 1996). In the late 1990s, educational

---

24 There have been significant changes of how the planning was implemented in recent years. One example is the “autonomous recruitment of students” by the universities (zizhu zhaosheng), of which the admission procedures are different from standardized college entrance exam.

25 Setting a target school enrollment rate is a common practice in Chinese educational reforms. For example, the 1993 *Outline of the Chinese Educational Reform and Development* announced that “in the 1990s …the illiteracy rate of the youth should be lower than 5 percent” (Ministry of Education, 1993); in 2001, the *Decision of the State Council on the Reform and Development of Basic Education* claimed that “in the Tenth Five-Year Plan the enrollment rate of senior high school should reach about 60 percent” (Ministry of Education, 2001).
policies began to seek to “actively develop both vocational and senior high school education”,
and “senior high school education should be universalized step by step in urban and more
developed areas” (Ministry of Education, 1998, 1999). Since then the enrollment rates for both
senior high school and vocational school have been quickly increasing. In 2002, the gross
enrollment rate (máo rúxuélǜ) of upper secondary education was 55.8% (Ministry of Education,
2004). In 2009, the figure climbed to 81.2% (Ministry of Education, 2011), suggesting that
around 80% of children aged 15-18 enrolled in upper secondary education.

The expansion at the college level (including junior college and university) was more
striking. In the Maoist era, colleges and universities were once closed during 1966-1969, and
only a small proportion of colleges were allowed to offer courses in the subsequent years (Unger
1982). In 1977, a merit-based College Entrance Exam was available for the first time, marking
the start of college expansion in the post-Mao period. The normalization of college education
began in the early 1980s. Since then full-time college education has become increasingly
available, mainly to senior high school graduates when they have passed the College Entrance
Exam. In 1995, the gross college enrollment rate was 6.5% (Ministry of Education, 1996). In
1998, the rate became 9.8% for individuals aged 18-22, indicating that for every ten people aged
18-22, there was about one enrolled college student (Ministry of Education, 2000). However,
large-scale increases in the college enrollment started in 1999. In that year, the Ministry of
Education released a policy document called Facing 21 Century: Acting Plans for Educational
Vitalization, aiming to enlarge college enrollment so that “the (gross) college enrollment rate
should reach about 11 percent by 2000 and should be close to 15 percent by 2010” (Ministry of
Education, 1998). Because of this policy, the Chinese government increased the target number of

---

26 In 1977, the College Entrance Exam was available to senior high school graduates from all the cohorts who did
not have the chance to enter college during the Cultural Revolution. In the subsequent two years, these people were
still allowed to take the College Entrance Exam if they were not admitted in 1977.
newly-admitted college students year by year through its national enrollment planning system, resulting in a larger and larger body of college students in subsequent years. The gross college enrollment rate reached 24.2% in 2009 and 35.4% in 2013, and was planned to be 40% by 2020 (Ministry of Education, 2011, 2014), much higher than anticipated in 1998.

To illustrate the trends of changing educational opportunities, Table 3.1 presents the national statistics on the transition rates by different levels of education in China’s reform period. As the figures indicate, the rate of transition to junior high school given the completion of primary school education declined substantially in the early-1980s and then quickly increased from 68.3 percent in 1981 to over 90 percent in 1996, and to 99.5 percent in 2008 (Column 1). The second and third columns of Table 3.1 present the rates of transition to upper secondary schools upon finishing compulsory education. They show that the expansion of educational opportunities at the upper secondary school level was contacted in the 1980s and was slow 1990s, but became relatively quicker since the late 1990s. According to the second column, the expansion was pronounced for the changing opportunities for senior high school education (the academic track). As the figures suggest, the rate of transition to senior high school given the completion of junior high school experienced an initial sharp drop from over 40 percent in 1978 to around 21 percent in the late 1980s, and remained pretty stable in the 1990s: in 1998, the rate was still lower than 23 percent. The only discernable expansion at the senior high school level started from 1999: the transition rate increased by 2.1 percent from the year 1998 to 1999, 1.2 times the 10 years’ accumulative increase from 1988 to 1998. The last decade since 1999 witnessed further quick expansion of senior high school opportunities, with the transition rate going from 24.9 percent in 1999 to 47.8 percent in 2010, or being increased by 1.9 times in the eleven years.
Table 3.1: Educational Expansion in China, National Statistics, 1977-2010

<table>
<thead>
<tr>
<th>Year</th>
<th>Transition to junior high school %</th>
<th>Transition to senior high school %</th>
<th>Transition to vocational school %</th>
<th>Transition to college %</th>
<th>Number of Newly-Admitted College Students (10 thousands)</th>
<th>Odds of transition rates, college vs. senior high</th>
</tr>
</thead>
<tbody>
<tr>
<td>1977</td>
<td>92.0&lt;sup&gt;a&lt;/sup&gt;</td>
<td>63.7</td>
<td>2.3</td>
<td>4.7</td>
<td>27.3</td>
<td>0.07&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>1978</td>
<td>87.7</td>
<td>40.9</td>
<td>2.6</td>
<td>5.9</td>
<td>40.1</td>
<td>0.14</td>
</tr>
<tr>
<td>1979</td>
<td>82.8</td>
<td>37.0</td>
<td>3.0</td>
<td>3.8</td>
<td>27.5</td>
<td>0.10</td>
</tr>
<tr>
<td>1980</td>
<td>75.5</td>
<td>39.7</td>
<td>8.0</td>
<td>4.6</td>
<td>28.1</td>
<td>0.11</td>
</tr>
<tr>
<td>1981</td>
<td>68.1</td>
<td>28.4</td>
<td>6.1</td>
<td>5.7</td>
<td>27.9</td>
<td>0.20</td>
</tr>
<tr>
<td>1982</td>
<td>65.9</td>
<td>27.1</td>
<td>8.2</td>
<td>10.1</td>
<td>31.5</td>
<td>0.37</td>
</tr>
<tr>
<td>1983</td>
<td>66.5</td>
<td>27.1</td>
<td>12.9</td>
<td>16.6</td>
<td>39.1</td>
<td>0.61</td>
</tr>
<tr>
<td>1984</td>
<td>65.3</td>
<td>27.6</td>
<td>15.6</td>
<td>25.0</td>
<td>47.5</td>
<td>0.91</td>
</tr>
<tr>
<td>1985</td>
<td>67.5</td>
<td>25.8</td>
<td>18.3</td>
<td>31.5</td>
<td>61.9</td>
<td>1.22</td>
</tr>
<tr>
<td>1986</td>
<td>68.8</td>
<td>24.3</td>
<td>17.1</td>
<td>25.5</td>
<td>57.2</td>
<td>1.05</td>
</tr>
<tr>
<td>1987</td>
<td>68.2</td>
<td>22.8</td>
<td>16.5</td>
<td>25.0</td>
<td>61.7</td>
<td>1.09</td>
</tr>
<tr>
<td>1988</td>
<td>69.4</td>
<td>21.1</td>
<td>17.0</td>
<td>26.7</td>
<td>67.0</td>
<td>1.27</td>
</tr>
<tr>
<td>1989</td>
<td>70.5</td>
<td>21.3</td>
<td>16.9</td>
<td>24.5</td>
<td>59.7</td>
<td>1.15</td>
</tr>
<tr>
<td>1990</td>
<td>73.5</td>
<td>22.5</td>
<td>17.7</td>
<td>26.1</td>
<td>60.9</td>
<td>1.16</td>
</tr>
<tr>
<td>1991</td>
<td>74.4</td>
<td>22.5</td>
<td>19.9</td>
<td>27.8</td>
<td>62.0</td>
<td>1.24</td>
</tr>
<tr>
<td>1992</td>
<td>78.2</td>
<td>21.3</td>
<td>21.8</td>
<td>33.3</td>
<td>75.4</td>
<td>1.57</td>
</tr>
<tr>
<td>1993</td>
<td>80.3</td>
<td>20.1</td>
<td>24.4</td>
<td>39.9</td>
<td>92.4</td>
<td>1.98</td>
</tr>
<tr>
<td>1994</td>
<td>85.1</td>
<td>21.1</td>
<td>25.8</td>
<td>43.0</td>
<td>90.0</td>
<td>2.04</td>
</tr>
<tr>
<td>1995</td>
<td>89.3</td>
<td>22.3</td>
<td>26.7</td>
<td>45.9</td>
<td>92.6</td>
<td>2.06</td>
</tr>
<tr>
<td>1996</td>
<td>91.0</td>
<td>22.1</td>
<td>26.7</td>
<td>47.1</td>
<td>96.6</td>
<td>2.14</td>
</tr>
<tr>
<td>1997</td>
<td>92.1</td>
<td>22.4</td>
<td>25.9</td>
<td>45.1</td>
<td>100.0</td>
<td>2.02</td>
</tr>
<tr>
<td>1998</td>
<td>92.6</td>
<td>22.8</td>
<td>24.3</td>
<td>43.1</td>
<td>108.4</td>
<td>1.89</td>
</tr>
<tr>
<td>1999</td>
<td>92.9</td>
<td>24.9</td>
<td>22.5</td>
<td>60.7</td>
<td>159.7</td>
<td>2.44</td>
</tr>
<tr>
<td>2000</td>
<td>93.6</td>
<td>29.4</td>
<td>19.6</td>
<td>73.2</td>
<td>220.6</td>
<td>2.49</td>
</tr>
<tr>
<td>2001</td>
<td>94.2</td>
<td>32.7</td>
<td>18.3</td>
<td>78.8</td>
<td>268.3</td>
<td>2.41</td>
</tr>
<tr>
<td>2002</td>
<td>95.8</td>
<td>36.0</td>
<td>19.8</td>
<td>83.5</td>
<td>320.5</td>
<td>2.32</td>
</tr>
<tr>
<td>2003</td>
<td>96.8</td>
<td>37.7</td>
<td>20.3</td>
<td>83.4</td>
<td>382.2</td>
<td>2.21</td>
</tr>
<tr>
<td>2004</td>
<td>97.3</td>
<td>39.7</td>
<td>20.9</td>
<td>81.8</td>
<td>447.3</td>
<td>2.06</td>
</tr>
<tr>
<td>2005</td>
<td>97.9</td>
<td>41.7</td>
<td>23.8</td>
<td>76.3</td>
<td>504.5</td>
<td>1.83</td>
</tr>
<tr>
<td>2006</td>
<td>99.7</td>
<td>42.2</td>
<td>27.8</td>
<td>75.1</td>
<td>546.1</td>
<td>1.78</td>
</tr>
<tr>
<td>2007</td>
<td>99.7</td>
<td>42.9</td>
<td>30.9</td>
<td>71.8</td>
<td>565.9</td>
<td>1.67</td>
</tr>
<tr>
<td>2008</td>
<td>99.5</td>
<td>44.9</td>
<td>32.1</td>
<td>72.7</td>
<td>607.7</td>
<td>1.62</td>
</tr>
<tr>
<td>2009</td>
<td>99.0</td>
<td>46.3</td>
<td>34.9</td>
<td>77.6</td>
<td>639.5</td>
<td>1.68</td>
</tr>
<tr>
<td>2010</td>
<td>98.6</td>
<td>47.8</td>
<td>34.1</td>
<td>83.3</td>
<td>661.8</td>
<td>1.74</td>
</tr>
</tbody>
</table>


<sup>a</sup> Figures in the first four columns are transition rates calculated from *China Statistical Yearbook* 1990, 2004-2011.

<sup>b</sup> Figures are the odds of transition rates to college vs. to senior high school.
Similarly, although the rate of transition to vocational upper secondary school (the vocational track) given compulsory education increased sharply in the 1980s from 2.6 percent in 1978 to around 20 percent in the early 1990s, the expansion of vocational school opportunity was mostly steady in the 1990s, with an average transition rate of 24.2 percent. In the first decade of the new century, the rate increased by 1.8 times, from 19.6 percent in 2000 to 34.9 percent in 2009. These figures suggest that the vocational track has been one important channel to absorb the increasing number of students with compulsory education.

College expansion was more dramatic. As figures in Column 4 show, the rate of transition to college upon completion of senior high school grew from 5.7 percent in 1981 to 33.3 percent in 1992; since 1993, the transition rate had increased less than 7 percent by 1998. Yet, starting from 1999, the rate has largely increased, jumping from 43.1 percent in 1998 to 60.7 percent in 1999, then to 83.4 percent in 2003, with an average increase rate by 48.3 percent per year. In the next four years, the rate decreased a little but remained as high as 72 percent in 2007, 1.67 times of that in 1998. In 2010, the rate had come back to around 83 percent. Column 5 further presents the number of newly-admitted college students each year. Again, it shows that during the twelve years from 1998 to 2010, college opportunities have been greatly shifted up. The number of newly-admitted college students has continuously increased from 1.08 million in 1998 to 6.61 million in 2010, i.e. by 6.12 times.

To interpret the figures more intuitively, Figure 3.2 summarizes the trends of educational expansion at various levels.
Figure 3.2: National Statistics of the Transition Rates, by Levels of Education

The plot shows that although the transition rate to junior high school increased in the 1980s, substantive increases only happened in the 1990s; after mid-1990s, junior high school education became almost universal. In contrast, the rate of transition to senior high school experienced a steadily decline in the 1980s, and was almost stable in the 1990s; after 1998, there was a quick increase in the conditional probability of progressing to senior high school. A similar trend holds for the vocational upper secondary school since the 1990s. When it comes to the college level, the transition rate kept increasing from the 1980s to the 1990s; after 1999, the increase in the transition rate was dramatic.

Because of these trends, China’s educational expansion has two notable features in the distribution of educational opportunities. First, compared to the 1980s, the 1990s witnessed a
contraction in the senior high school opportunity. As the transition rates to senior high school were declining in the 1980s, it was more difficult for a junior high school graduate to enter senior high school in the 1990s than in the 1980s. Moreover, in the 1990s, while there was an increasing supply of students with compulsory education, the chances of entering senior high school were almost the same in that decade (the transition rates to senior high school were mostly stable in the 1990s; see the first scenario in Chapter 2).

Secondly, the probability of transition to college was higher than the transition to senior high school since the mid-1980s, and the gap between the two probabilities has increased over time, particularly in the first several years after 1999 (this is the case of $E_k < E_{k+1}$ discussed in Chapter 2). As shown in Figure 3.1, starting from 1999, the transition rates to senior high school and to college have both been increasing. However, the increase was more gradual for the former, as compared to the “sudden jump” of the latter. In this sense, the Chinese educational system has a “bottleneck” structure since the mid-1980s, with senior high school education being the “neck” for continuing post-compulsory education (see also Hannum et al 2011). The last column of Table 3.1 calculates the odds of the transition rates to college over to senior high school, which measure the gap between expansions at the two levels. It shows that while the odds have remained larger than 1 since 1985, only in the first several years after 1999 did they reach the highest points over the last thirty years (i.e. 2.49 in 2000). Given the high transition rate to college since 1999, China’s “bottleneck” structure seemed to be more constrained for transitions to post-compulsory education after the 1999 expansion.  

---

27 In Chapter 2, I have examined several scenarios of educational expansion at level $k$ and level $k+1$, which could be used to understand China’s “bottleneck” structure. More detailed discussion of how the case of reforming China fits in the general theory in Chapter 2 will be presented in Chapter 4.
Financing Education

Along with the changing educational opportunities at all levels, educational reforms have also changed how education is financed. The finance for education mainly comes from four sources: budgetary educational appropriation (yusuannei bokuan), educational surcharges from the central and local governments, educational revenues from school-run enterprises and social services, and school undertaking revenues (shiye shouru) including tuition and fees (Hu and Shen 2007; Shen and Du 2009; Wang 2012). Table 3.2 presents the national statistics on the components of educational finance for different levels of education from 1995 to 2010.

As the first two columns show, budgetary educational appropriation from the governments has been the most important financial source for compulsory education. At the junior high school level, its percentage of the total educational revenue increased from 53.0 percent in 1995 to 71.1 percent in 2005. In 2006, to realize the goal of free nine-year education, the amended Compulsory Education Law required the finance of compulsory education be guaranteed by the central and local governments (Ministry of Education, 2006). School fees for compulsory education were beginning to be eliminated in western rural areas in that year, and were completely eradicated in all rural areas in 2007. Similar measures were extended to urban areas in 2008 (Xinhua Net, 31/07/2008). As we can see from the second column, starting from 2006, the percentage of budgetary educational appropriation jumped from 77.0 percent to over 90 percent in 2010, suggesting that fiscal budgetary appropriation from the governments has become the single most significant component for financing compulsory education in recent years.
<table>
<thead>
<tr>
<th>Year</th>
<th>Junior High School</th>
<th>Senior High School</th>
<th>Vocational School</th>
<th>College</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Overall (in billions)</td>
<td>Budgetary appropriation %</td>
<td>Overall (in billions)</td>
<td>Budgetary appropriation %</td>
</tr>
<tr>
<td>1995</td>
<td>39.3</td>
<td>53.0</td>
<td>13.7</td>
<td>51.5</td>
</tr>
<tr>
<td>1996</td>
<td>47.9</td>
<td>52.6</td>
<td>16.9</td>
<td>49.1</td>
</tr>
<tr>
<td>1997</td>
<td>51.7</td>
<td>54.3</td>
<td>20.0</td>
<td>47.2</td>
</tr>
<tr>
<td>1998</td>
<td>47.2</td>
<td>55.2</td>
<td>32.9</td>
<td>46.3</td>
</tr>
<tr>
<td>1999</td>
<td>51.1</td>
<td>57.7</td>
<td>38.5</td>
<td>45.8</td>
</tr>
<tr>
<td>2000</td>
<td>57.0</td>
<td>59.9</td>
<td>45.6</td>
<td>44.3</td>
</tr>
<tr>
<td>2001</td>
<td>68.1</td>
<td>63.8</td>
<td>57.6</td>
<td>45.4</td>
</tr>
<tr>
<td>2002</td>
<td>79.4</td>
<td>68.3</td>
<td>71.8</td>
<td>43.7</td>
</tr>
<tr>
<td>2003</td>
<td>88.5</td>
<td>69.0</td>
<td>82.6</td>
<td>43.3</td>
</tr>
<tr>
<td>2004</td>
<td>101.8</td>
<td>70.3</td>
<td>98.0</td>
<td>44.4</td>
</tr>
<tr>
<td>2005</td>
<td>119.4</td>
<td>71.1</td>
<td>114.7</td>
<td>45.0</td>
</tr>
<tr>
<td>2006</td>
<td>134.3</td>
<td>77.0</td>
<td>129.5</td>
<td>48.1</td>
</tr>
<tr>
<td>2007</td>
<td>188.8</td>
<td>84.0</td>
<td>127.0</td>
<td>53.8</td>
</tr>
<tr>
<td>2008</td>
<td>236.1</td>
<td>87.5</td>
<td>144.8</td>
<td>57.2</td>
</tr>
<tr>
<td>2009</td>
<td>280.7</td>
<td>89.8</td>
<td>161.4</td>
<td>60.2</td>
</tr>
<tr>
<td>2010</td>
<td>321.0</td>
<td>90.5</td>
<td>182.8</td>
<td>63.7</td>
</tr>
</tbody>
</table>


Note: The statistics of senior high school revenues include those for “combined junior and senior high schools” (*wanquan zhongxue*) in 1998-2006 (the only statistics available). The statistics of vocational school revenues are the sum of those for technical schools, associate schools and vocational schools. The statistics of college refer to the revenues for full-time universities (*benke*) and junior colleges (*dazhuan*).
In contrast, upper secondary schools are fee-bearing. Compared to junior high school, the second panel shows that although the total revenue for senior high school was increasing in the 15 years, educational input from the governments was limited. The percentage of budgetary appropriation was declining from 51.5 percent in 1995 to 43.7 percent in 2003, and then increased gradually to 63.7 percent in 2010. In the meanwhile, educational revenue from school tuition and fees increased fast, from 13.5 percent in 1995 to 23.5 percent in 2007, i.e. by 74 percent in twelve years. A similar trend holds for vocational education. The percentage of tuition and fees of the total revenue remained as high as 30 percent in 2006, about three fifths of the budgetary appropriation for vocational schools. These figures indicate that tuition and fees have been an increasingly important channel for financing upper secondary education.

The last panel of Table 3.2 shows the changes of educational finance for college education. Total revenue for college grew by 18.5 times in the 15 years, but the governmental input on higher education was declining in most years of the period. In 1995, budgetary appropriation from the governments accounted for about 70 percent of the total college fund, in 1999 the figure dropped to 49.4 percent. At the same time, the proportion of tuition and fees as financial input was increased, from 13.6 percent in 1995 to 17.2 percent in 1999. There figures reflect changes in the college financing policies in the reform periods. In the 1980s and early 1990s, the Chinese college education was mainly funded by the governments, being almost free.\(^{28}\) Starting from 1994, a small amount of money was charged to all the college attendants as their tuition and fees. By 1997, self-payment policies were fully implemented (Hannum et al 2012). In 1999, along

\(^{28}\) During 1986 to 1994, there was a “dual track” system (shuangui zhi) in the admission process of higher education: no tuition or fee was charged for those who successfully passed the College Entrance Exam (jihua sheng); in the meantime, colleges could also admit a very small amount of students (zifei sheng) whose College Entrance Exam score was lower than the regular admission criteria, provided that they could pay for the score difference. Starting from 1994, the Ministry of Education closed the admission for zifei sheng (Ministry of Education, 1994). In addition, since the 1990s, following the Soviet model, not only free, but students received stipends as future income.
with the intention to make tertiary education more available to the youth, the central government raised the tuition and fees for college so as to “marketize higher education” (Dai and Mo 2004; Liu 2004, 2005; Min 2007). Since then, the percentage of governmental input on college education of the total revenue has been declining in subsequent years, and the percentage of tuition and fees as a means for financing college has been increasing. In 2005, tuition and fees accounted for about one third (33.8%) of the total educational fund, and over 80 percent of the budgetary appropriation provided by the governments. Although in recent years governmental expenditure on higher education was increasing, tuition and fees were still about one fourth of the total college input in 2010, and 43.5 percent (24.0/55.1) of the governmental financial allocations.

Table 3.3 further calculates individual educational costs for senior high school and college education based on national statistics. The first panel presents tuition and fees per student in senior high and college. The figures suggest that individual educational cost was increasing for both levels in the past 15 years, but the gap has widened as the cost for college increased faster than that for senior high school. Particularly, along with the massive educational expansion since 1999, individual tuition and fees rose significantly at both levels. In 1998, the average tuition and fees for senior high school was 393 yuan; in 2010, the figure became 1418 yuan. After adjusted by the consumer price index (CPI), the average tuition for senior high school has increased by 2.95 times in the twelve years, with an average increase rate of 24.6 percent. Average college tuition and fees also rose fast. In 1998, the average tuition and fees for college students was 1974 yuan; in 1999 the figure rose to 2769 yuan, with a 40.3 percent increase. It continued to increase in subsequent years. In 2002, the CPI-adjusted average tuition and fees became 2.16 times of that in four years ago. In 2010, the average college tuition and fees reached
Table 3.3: Educational Costs for Senior High School and College, National Statistics, 1995-2010

<table>
<thead>
<tr>
<th>Year</th>
<th>Average tuition and fees (yuan)</th>
<th>GDP and household income per capita (yuan)</th>
<th>% Average tuition and fees: Senior High School</th>
<th>% Average tuition and fees: College</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Senior high school</td>
<td>College</td>
<td>GDP per capita</td>
<td>Urban disposable income</td>
</tr>
<tr>
<td>1995</td>
<td>267</td>
<td>1001</td>
<td>5046</td>
<td>4283.0</td>
</tr>
<tr>
<td>1996</td>
<td>330</td>
<td>1315</td>
<td>5846</td>
<td>4838.9</td>
</tr>
<tr>
<td>1997</td>
<td>393</td>
<td>1589</td>
<td>6420</td>
<td>5160.3</td>
</tr>
<tr>
<td>1998</td>
<td>415</td>
<td>1974</td>
<td>6796</td>
<td>5425.1</td>
</tr>
<tr>
<td>1999</td>
<td>490</td>
<td>2769</td>
<td>7159</td>
<td>5854.0</td>
</tr>
<tr>
<td>2000</td>
<td>641</td>
<td>3550</td>
<td>7858</td>
<td>6280.0</td>
</tr>
<tr>
<td>2001</td>
<td>758</td>
<td>3895</td>
<td>8622</td>
<td>6859.6</td>
</tr>
<tr>
<td>2002</td>
<td>860</td>
<td>4224</td>
<td>9398</td>
<td>7702.8</td>
</tr>
<tr>
<td>2003</td>
<td>948</td>
<td>4419</td>
<td>10542</td>
<td>8472.2</td>
</tr>
<tr>
<td>2004</td>
<td>1020</td>
<td>4785</td>
<td>12336</td>
<td>9421.6</td>
</tr>
<tr>
<td>2005</td>
<td>1118</td>
<td>5112</td>
<td>14185</td>
<td>10493.0</td>
</tr>
<tr>
<td>2006</td>
<td>1114</td>
<td>5075</td>
<td>16500</td>
<td>11759.5</td>
</tr>
<tr>
<td>2007</td>
<td>1195</td>
<td>5779</td>
<td>20169</td>
<td>13785.8</td>
</tr>
<tr>
<td>2008</td>
<td>1190</td>
<td>6113</td>
<td>23708</td>
<td>15780.8</td>
</tr>
<tr>
<td>2009</td>
<td>1338</td>
<td>6240</td>
<td>25608</td>
<td>17174.7</td>
</tr>
<tr>
<td>2010</td>
<td>1418</td>
<td>6383</td>
<td>29992</td>
<td>19109.4</td>
</tr>
</tbody>
</table>


*Figures are calculated from *China Educational Finance Statistical Yearbook, 1996-2011.*

*Figures are the numbers in the first panel divided by those in second panel, respectively.*
Note: The dotted lines are adjusted by the 1995 Consumer Price Index (CPI).

Figure 3.3: Tuition and Fees per Student for Senior High School and College (in RMB), National Statistics, 1995-2010

up to 6383 yuan, 3.1 times (2.64 times if adjusted by CPI) of the average individual cost in 1998.

Figure 3.3 summarizes the trend of increasing schooling cost since the late 1990s. It shows that due to the educational policy in 1999, educational cost has been dramatically shifted up both for senior high school and for college.

The third and fourth panels of Table 3.3 additionally compute the rates of the average tuition and fees over GDP per capita and mean household income for the urban and rural families, respectively. Figures in the third panel show that the average schooling fees for high school education were about 5 to 10 percent of GDP per capita over the 15 years. The rate increased from 5.3 percent in 1995 to 9.2 percent in 2002, and then went down to 4.7 percent in 2010. Before 1999, urban families on average spent less than 8 percent of household disposable income
on senior high school charges. The figure went up to around 11 percent in the next few years. The rate was much higher for rural families. In the late 1990s, the average schooling cost for senior high school accounted for about 19 percent of the mean rural household income. Since then the rate has been doubled in 2004 and remained around 30 percent in 2007. In 2010, rural families still needed to spend an average rate of 24 percent of their household income for children’s senior high school education.

The sharp rise of college tuition and fees since the late 1990s has imposed a more substantial burden on both urban and rural families. The last panel in Table 3.3 reveals that the average tuition and fees for college education accounted for less than 30 percent of GDP per capita in 1998, but the rate jumped to 38.7 percent in 1999 and to 45.2 percent in 2000. In the next decade the rate was declining because of China’s rising GDP per capita, but remained about 30 percent of GDP per capital in 2006 and over one fifths in 2010. Because of the 1999 college expansion, the rate of average college cost over the mean urban disposable household income increased quickly, from 36.4 percent in 1998 to 47.3 percent in 1999. In the next few years after the expansion, the rate was over 50 percent, suggesting that urban families on average needed to put more than half of the household disposable income on children’s college education. For rural families, the average college cost was even higher than the mean annual net household income after the 1999 college expansion. In 1998, rural families on average had to spend 91.3 percent of household net income to pay for children’s college education. In 2002, the mean college tuition and fees was 1.71 times of an average rural family’s net income. Although the figure was declining in the last decade, paying for children’s college has been one of the largest expenses for rural families (Shen and Du 2009). In 2010, the average tuition and fees was still 1.08 times of the mean rural household net income.
Figure 3.4 summarizes the rates in the last panel of Table 3.3.

Clearly, it shows that affording for college education has been more challenging for rural families than for urban families. The gap between the rural and urban rates became larger after the 1999 college expansion, suggesting that rural children now found it more difficult to pay for college as compared to their urban peers (this means a lower $\varphi$ parameter in the language of Chapter 2). Therefore, the rising educational cost since 1999 has put rural children into a more disadvantaged position than before (Li 2006, Shen 2011).

THE CONSEQUENCES OF EDUCATIONAL EXPANSION

My introduction in the former section shows that China’s educational expansion in the reform periods has introduced not only more educational opportunities at all levels, but also
rising tuition and fees for the finance of education. In Chapter 2, I conceptualize the changing educational opportunities as a “structural change” and the changing educational cost as an “institutional change” (Figure 2.2, Chapter 2). In the following, I discuss how the structural and institutional changes introduced by educational expansion are relevant for understanding education-based inequality and mobility in reforming China.

**Educational Attainment**

As mentioned before, one special characteristic of China’s educational system in the reform periods is its national enrollment planning system. The central government usually sets target school enrollment rates by its educational policies, and accordingly makes national enrollment plans on a yearly basis. Based on a complicated but intransparent formula, the planned national school enrollments are then distributed to each province and provincial level educational authorities finalize the distribution of the enrollment “quota” through school entrance examinations (Wang et al 2014). Since the school enrollment is planned ahead and the “quota” of admission is generally fixed within each province, there are certain cutoff points of the required exam scores for school entry each year, which are determined by provincial level (and county level for senior high school admission) educational authorities. Therefore, national enrollment planning is realized through the control of the cutoff points of the school entrance exam scores (*luqu fenshuxian*) in each province. A larger national school enrollment usually implies a lower requirement of academic performance (or academic ability) by schools, as schools now need to incorporate more students based on the entrance exam scores, which is market demand-driven.

In this regard, the dramatic college expansion since 1999 has significantly lowered the
university requirement of students’ academic performance on the College Entrance Exam. Students whose academic performance was not good enough to enter college before the expansion now could enjoy the college opportunity provided by the expansion. As more and more students who would be considered as less qualified before the expansion rushed into higher education in a short period of time, universities became less selective, and the value of college credentials tended to decline in the labor market (Liu 2005).

However, the increased college opportunity after the expansion was not evenly distributed among children with different social origins (Li 2; Wu and Zhang 2010). The unequal distribution of college opportunities may even be strengthened by the structural and institutional changes in educational system along the process of educational expansion. As Figure 3.1 indicates, in the Chinese educational system most students have to complete the academic track upper secondary education (senior high school) to apply for college. Nevertheless, the senior high school opportunity was rather limited. Only less than 23 percent of junior high school graduates were able to enter senior high school before 1999; in 2010, the percentage was still less than 50 percent (Table 3.1). Since senior high school has been the “bottleneck” for continuing post-compulsory education, the selection of students who could have passed through the “bottleneck” is directly linked to those who could go to college, particularly given the high transition rate to college after the expansion. As a result, along with the declining requirement of academic performance by colleges, a higher percentage in the senior high school implies a larger likelihood to enjoy the benefits of college expansion.

Based on the available national statistics, Table 3.4 presents the distribution of educational opportunities at the senior high school level by different localities of junior and senior high schools.
Table 3.4: Transition Rates to Senior High School by School Localities, National Statistics, 1985-2003

<table>
<thead>
<tr>
<th>Year</th>
<th>Urban cities and towns</th>
<th>Counties and towns</th>
<th>Villages</th>
<th>Counties, towns and villages</th>
<th>Odds of transition rates, urban cities vs. others</th>
</tr>
</thead>
<tbody>
<tr>
<td>1985</td>
<td>40.3 (^a)</td>
<td>58.7</td>
<td>11.6</td>
<td>22.3</td>
<td>1.81 (^b)</td>
</tr>
<tr>
<td>1986</td>
<td>40.2</td>
<td>55.6</td>
<td>10.2</td>
<td>20.9</td>
<td>1.92</td>
</tr>
<tr>
<td>1987</td>
<td>40.2</td>
<td>53.8</td>
<td>9.7</td>
<td>19.5</td>
<td>2.06</td>
</tr>
<tr>
<td>1988</td>
<td>37.7</td>
<td>51.1</td>
<td>8.6</td>
<td>17.8</td>
<td>2.11</td>
</tr>
<tr>
<td>1989</td>
<td>38.4</td>
<td>51.3</td>
<td>8.3</td>
<td>17.9</td>
<td>2.14</td>
</tr>
<tr>
<td>1990</td>
<td>40.4</td>
<td>54.2</td>
<td>8.7</td>
<td>19.0</td>
<td>2.13</td>
</tr>
<tr>
<td>1991</td>
<td>40.1</td>
<td>52.7</td>
<td>8.1</td>
<td>18.9</td>
<td>2.12</td>
</tr>
<tr>
<td>1992</td>
<td>38.5</td>
<td>46.9</td>
<td>7.2</td>
<td>17.8</td>
<td>2.16</td>
</tr>
<tr>
<td>1993</td>
<td>37.3</td>
<td>40.4</td>
<td>6.4</td>
<td>16.5</td>
<td>2.27</td>
</tr>
<tr>
<td>1994</td>
<td>39.1</td>
<td>41.5</td>
<td>6.3</td>
<td>17.0</td>
<td>2.30</td>
</tr>
<tr>
<td>1995</td>
<td>41.6</td>
<td>42.6</td>
<td>6.5</td>
<td>17.9</td>
<td>2.33</td>
</tr>
<tr>
<td>1996</td>
<td>44.1</td>
<td>40.8</td>
<td>6.3</td>
<td>17.3</td>
<td>2.55</td>
</tr>
<tr>
<td>1997</td>
<td>44.9</td>
<td>39.9</td>
<td>6.1</td>
<td>17.3</td>
<td>2.60</td>
</tr>
<tr>
<td>1998</td>
<td>47.2</td>
<td>40.2</td>
<td>6.0</td>
<td>17.3</td>
<td>2.74</td>
</tr>
<tr>
<td>1999</td>
<td>55.4</td>
<td>44.6</td>
<td>6.2</td>
<td>18.6</td>
<td>2.98</td>
</tr>
<tr>
<td>2000</td>
<td>66.7</td>
<td>52.8</td>
<td>7.1</td>
<td>22.1</td>
<td>3.02</td>
</tr>
<tr>
<td>2001</td>
<td>70.6</td>
<td>49.7</td>
<td>8.0</td>
<td>25.5</td>
<td>2.77</td>
</tr>
<tr>
<td>2002</td>
<td>74.2</td>
<td>53.1</td>
<td>8.8</td>
<td>28.3</td>
<td>2.63</td>
</tr>
<tr>
<td>2003</td>
<td>77.4</td>
<td>56.5</td>
<td>8.7</td>
<td>29.3</td>
<td>2.64</td>
</tr>
</tbody>
</table>


Note: For each locality of schools, the yearly transition rate is the number of newly-admitted senior high school students divided by the number of newly-graduated junior high school students. National statistics are only available in 1985-2003.

\(^a\) Figures in the first four columns are transition rates calculated from *China Statistical Yearbook* 1986-2004.

\(^b\) Figures are the odds of transition rates in the first column divided by those in the fourth column.

The transition rates in the first column of Table 3.4 are computed using the number of newly-admitted senior high school students and the number of newly-graduated junior high school students in urban cities each year. Similarly, figures in the next two columns refer to the yearly transition rate for students who go to junior and senior high schools in counties and towns (*xianzhen*), and villages, respectively. Figures in the fourth column are the total number of newly-admitted senior high students in counties, towns and villages divided by the total number of newly-graduated junior high school students in the same localities. Since most rural students
go to junior and senior high schools in counties, towns, and villages (Zhang 2010), the transition rate presented in the fourth column is a rough measure of the likelihood of progressing to senior high school for rural children upon finishing junior high school.

Comparing the transition rates in the first and fourth column indicates that the probability of going to senior high school differed significantly for urban and rural children. Not only did children in cities have a higher chance to go to senior high than their counterparts in counties and rural areas, but their schooling opportunity also grew faster in most years. In 1998, for instance, the transition rate was 47.2 percent for children in cities and 17.3 percent for children in counties and rural areas. In the next four years, the figure rose up to 77.4 percent and 29.3 percent, respectively. The last column of Table 3.4 approximates the odds of the transition rate to senior high school for urban and rural children. Clearly, the urban-rural gap (roughly measured by the odds) in the transition rate had been enlarging since 1985 and reached its peak in 2000 (odds=3.02). The pattern is more intuitively illustrated in Figure 3.5. As the figure shows, after 1998, while the overall rate of transition to senior high school increased fast, the distribution of senior high school opportunity by the localities of schools became more unequal than before. Children in cities had much higher likelihood of entering senior high school than those who received secondary education in counties and rural areas after the 1999 educational expansion.

The changing distribution of schooling opportunities at the senior high school level can be transmitted into the college attainment process. If, as Figure 3.5 suggests, certain social groups enjoy advantages for going to senior high school, they are then inclined to accumulate their advantages in obtaining college opportunities simply because of the high probability of entering college after the expansion. Therefore, China’s “bottleneck” structure in the educational
system tends to enlarge the unequal distribution of college opportunities after massive school expansion.

The rise of educational costs for senior high school and college education further deepens the opportunity gap in college attainment among different social groups. As illustrated in Table 3.3, paying for children’s education has been a substantial financial challenge for the Chinese families, and the burden was much more heavy on disadvantaged social groups (rural families in this case). A survey conducted in 2005 suggested that many families fell into poverty by having to pay their children’s tuition and surcharges in senior high school and college (Liu and Yang 2007). The proportion of rural families that found it “difficult” or “very difficult” to pay these fees was at least over 39 percent. Along this line, one remarkable result is that children
from disadvantaged backgrounds, especially those from poor rural areas, had lower educational aspiration than before (Liu 2005). In 2009, over 10 thousand senior high school graduates in Chongqing voluntarily opted out the College Entrance Exam, most of whom were rural children (Xinhua Net, 01/04/2009). Recent observations suggest that the idea of “studying is useless” (dushu wuyong lun) has emerged again in rural areas (Li et al 2012). Some rural parents even refused to let the children go to college even though they were admitted (People’s Daily Online, 05/09/2013). In contrast, advantaged groups were affected less by the rising educational cost since the late 1990s. Children with advantaged backgrounds now had a better chance to enjoy the benefit of their family resources. Not only could they pay for education, but they could also invested more on education-related materials, activities and programs than before. In certain circumstances, there was even room for advantaged families to use money or privilege to help their children to receive higher education. In a word, the institutional change in the finance of education tends to shift up the opportunity gap between the advantaged and disadvantaged groups.

Therefore, the direct effects of educational expansion on educational attainment since 1999 are two-fold. On the one hand, the increasing educational opportunity (structural change) has made the Chinese “bottleneck” structure more constricted for children with disadvantaged backgrounds. When college opportunity was not widely available and competition for college was high, the College Entrance Exam ensured that only qualified students could enter college. Under this situation, those less qualified students from advantaged groups, though they passed

---

29 Other than the rising educational cost for college, the devaluation of college credentials in terms of income, employment and prestige, etc. could be another reason for lower educational aspiration (Li et al 2012). See also the discussion in footnote 6, Chapter 2.

30 Since the 1999 reform, some universities/colleges were allowed to independently use 5% of their enrollment quota, which opened the door for corruption in the admission process (People’s Daily Online, 05/04/2014).
the bottleneck constraint, were selected out of the competition for college. After the 1999 expansion, children who came from advantaged groups that could not have entered college under the selection criteria before now had the opportunity to receive higher education. At the same time, since senior high school education was still not widely available, the “bottleneck” barrier continued to prevent many children with disadvantaged backgrounds from further education. This means that only certain groups of children with disadvantaged family backgrounds could enjoy the fruit of college expansion.

On the other hand, the rise of educational cost (institutional change) widened the differences in the affordability of education between the advantaged and disadvantaged groups (see also Archibald and Feldman 2008). Children from poor families who viewed the educational cost as too high may decide to quit in earlier transitions, for example, by not continuing to senior high school or by not taking the College Entrance Exam. The rising educational cost was less consequential for the advantaged groups for college attainment.

The Changing Volume and Composition of Educated Workers

As established in Chapter 2 (Figure 2.1), educational expansion changes the social selection process of who can go to schools. Not only does the quantity of admitted students increase after the expansion, but the distributions of individual characteristics such as ability and social origin between the pre- and post-expansion students also differed. Therefore, the labor market entry cohorts before and after educational expansion were different in terms of both the volume and the composition of skilled labor. Without considering the differences between the pre- and post-expansion schooling cohorts, we underestimate the effects of educational expansion on the labor market returns to education (Carneiro and Lee 2007; Juhn et al 2005).
In the Chinese context, as discussed in the previous section, more students were able to enter college since the 1999 college expansion, and the average academic ability of college students declined. Particularly, because of the “bottleneck” structure, those who now enjoy schooling opportunities but would not have been qualified before the expansion but disproportionately came from better-off families (see Figure 3.5). Moreover, the rising educational costs tend to further strengthen the association between social origin and education. Together, in a short period of time since 1999, China’s educational expansion has produced a cohort of college students of larger quantity, lower quality, and higher origin-education association than before.

**Sources:** *China Statistical Yearbook* 2011; *China Labor Statistical Yearbook*, 2011.

**Figure 3.6: The Increasing Supply of College-educated Workers, National Statistics, 1977-2010**

To illustrate these changes, Figure 3.6 graphs the trend of increasing supply of college
graduates in the past three decades. As it shows, the number of newly-graduated college students grew exponentially after the 1999 college expansion. According to national statistics (National Bureau of Statistics 2011), in 2002, 1.34 million students graduated from college; in 2003, four years (the normal length of college education) after the 1999 expansion, the number of newly-graduated college students jumped to 1.87 million. In 2010, the number of newly-graduated college students was 5.75 million, 4.3 times of that in 2002. The dramatic increase of the supply of college students in the last decade is also reflected in the composition of urban total employment, as college graduates typically find jobs in urban areas. In 2002, newly-graduated college students accounted for 0.53 percent of the total urban employees; eight years later, the percentage shifted to 1.66 percent.

Table 3.5: Labor Market Composition by Education in China, National Statistics, 1997-2010

<table>
<thead>
<tr>
<th>Year</th>
<th>Level of Education %</th>
<th>Age 20-24</th>
<th>Age 25-29</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Illiterate</td>
<td>Primary</td>
<td>Junior high</td>
</tr>
<tr>
<td>1997</td>
<td>11.6</td>
<td>34.8</td>
<td>37.9</td>
</tr>
<tr>
<td>1998</td>
<td>11.5</td>
<td>34.2</td>
<td>38.9</td>
</tr>
<tr>
<td>1999</td>
<td>11.0</td>
<td>33.3</td>
<td>39.9</td>
</tr>
<tr>
<td>2000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2001</td>
<td>7.8</td>
<td>30.9</td>
<td>42.3</td>
</tr>
<tr>
<td>2002</td>
<td>7.8</td>
<td>30.0</td>
<td>43.2</td>
</tr>
<tr>
<td>2003</td>
<td>7.1</td>
<td>28.7</td>
<td>43.7</td>
</tr>
<tr>
<td>2004</td>
<td>6.2</td>
<td>27.4</td>
<td>45.8</td>
</tr>
<tr>
<td>2005</td>
<td>7.8</td>
<td>29.2</td>
<td>44.1</td>
</tr>
<tr>
<td>2006</td>
<td>6.7</td>
<td>29.9</td>
<td>44.9</td>
</tr>
<tr>
<td>2007</td>
<td>6.0</td>
<td>28.3</td>
<td>46.9</td>
</tr>
<tr>
<td>2008</td>
<td>5.3</td>
<td>27.4</td>
<td>47.7</td>
</tr>
<tr>
<td>2009</td>
<td>4.8</td>
<td>26.3</td>
<td>48.7</td>
</tr>
<tr>
<td>2010</td>
<td>3.4</td>
<td>23.9</td>
<td>48.8</td>
</tr>
</tbody>
</table>

Table 3.5 further presents the changing composition of educational groups in the labor market, using currently available national statistics. Figures in the first panel suggest that educational expansion since the late 1990s has indeed supplied more skilled workers to the labor market. In 1999, only 3.8 percent of the Chinese workers were college-educated; in 2010, 10.1 percent had received college education. Meanwhile, the percentage of workers with primary education or below decreased substantially, but the percentage of employees who received only upper secondary education increased only a little, from 11.9 percent in 1999 to 13.9 percent in 2010.

The next two panels in Table 3.5 present the distribution of educational groups for the young cohorts (aged 20-29). To save space, only the percentages of workers with at least upper secondary education are reported. Since the young cohorts were directly affected by the 1999 educational expansion, the proportion of worker with college education increased quickly. In 1997, only 3.3 percent of the 20-24 years old working population was college-educated; in 2010, 12.6 percent has gone to college. A similar trend holds for those aged 25-29: the percentage of college graduates increased from 4.7 percent in 1997 to 20.7 percent in 2010. The share of workers with upper secondary education also increased for the young cohorts.

The increasing share of worker with higher education shifts the composition of educational groups in the labor market. To indicate the trend of the changing compositions more intuitively, Figure 3.7 plots the upper and lower percentiles for those with college and upper secondary education (aged 25-29).
Note: The border lines of the area in maroon refer to the upper- and lower-limits of the percentiles for upper secondary school graduates.

Figure 3.7: The Changing Composition of Educational Groups for the Young Cohorts (Aged 25-29), National Statistics, 1997-2010

Statistics in the last panel of Table 3.5 show that in 1997, college graduates made up the upper 4.7 percentiles of the educational distribution for those aged 25-29, while upper secondary school graduates were between the 83rd (100-4.7-12.6) and 95th percentiles. By 2010, college graduates made up the upper 20.7 percentiles, while upper secondary school graduates were between the 63rd and 79th percentiles. Reflected in Figure 3.7, we see that the lower-limit percentile of upper secondary school graduates in 1997 was close to that of college graduates in 2010 (82.7th vs. 79.3th). Therefore, the ability of a certain proportion of college graduates in 2010 was more comparable to that of upper secondary school graduates in 1997 than to their college
graduate counterparts. As the share of college graduates grew, the figure also suggests that educational expansion has increased the (ability) heterogeneity among the young college graduates.

**Labor Market Returns to Education**

Since the demographic compositions of the labor market entry cohorts differ before and after educational expansion, the cohorts face different reward structures in the labor market. On the one hand, the sharp increase of the supply of college graduates within a short period tends to devalue the college credentials, net of other changes in the labor market (Carneiro and Lee 2001, 2007; Yang 2006). On the other hand, as the average cohort quality of college graduates declines after the expansion, whether viewing education as a human capital indictor (Rosen 1977) or a signaling device (Stiglitz 1975), both suggest that the mean amount of human capital and productivity indicated by the college credential is lower than before. Therefore, other things being equal, the linkage between college education and labor market outcomes tends to weaken after educational expansion, and the labor market returns to college tend to decline.

However, as suggested in Figure 3.7, the (ability) heterogeneity among college graduates has increased after educational expansion. Particularly, a certain proportion of college graduates are only comparable to those senior high school graduates before the expansion. Given China’s “bottleneck” structure and the rising educational cost, those lower-quality college students come disproportionately from advantaged social groups (see Figure 3.4 and 3.5), and the association between origin and college education tend to increase. According to the formal theorizing in Chapter 2, a “compositional effect” (Breen and Jonsson 2007) is thus evident in the Chinese

---

31 The assumption is that those in the same place in their cohort’s educational attainment distribution have about the same level of ability (Kurtzon 2012; Rosenbaum 1999).
context: that the larger origin-education association for college graduates and the larger share of college graduates in the labor market tend to predict a higher association between labor market outcomes and social origin for those with college education.

Taken together, China’s educational expansion posts two implications for understanding the linkage between education and labor market outcomes. First, it decreases the overall labor market returns to college. Second, it shifts up the effect of social origin on the labor market returns to college. 32

CONCLUSION

In this chapter, I illustrate the case of reforming China’s educational expansion so as to provide an empirical testing case for the general formal theory proposed in Chapter 2. My illustration of the case shows that China’s massive educational expansion since 1999 is characterized by an increasingly constrained “bottleneck” structure and rising educational costs for senior high school and college education. These structural and institutional changes under educational expansion have affected not only how educational opportunity is allocated, but also how education is rewarded in the labor market. Specifically, based on the theoretical framework in Chapter 2, I hypothesize that educational expansion would enlarge educational inequality at the college level, decrease the labor market returns to college education, and strengthen the effect of social origin on the labor market returns to college.

In the subsequent chapters, I will focus on the empirical patterns of the consequences of educational expansion in reforming China. Specifically, I will investigate how different social

32 However, China’s educational expansion happened along with its market transition. Market transition theory (Nee 1989; Nee and Matthews 1996) suggests that education will be better rewarded during the transition. In Chapter 1 I made a conceptual distinction between the effects of market transition and of educational expansion on educational returns. More discussion on the two sources of effects will be provided in Chapter 6.
groups vary in the opportunities to obtain education (Chapter 4) and first occupation (Chapter 5), and how education is rewarded in the labor market (Chapter 6). I will mainly address two sets of social backgrounds in the Chinese context: the class differentials and the urban-rural divide. Through this exercise, I aim to provide not only a vivid account of China’s education-based social inequality and mobility patterns, but also a demonstration of the general theory proposed in Chapter 2.
Chapter 4


INTRODUCTION

Scholars have long been interested in studying the trend of educational inequality in the context of educational expansion (Featherman and Hauser 1978; Mare 1980). In most industrial countries, researchers found that the effect of social origins on educational attainment is largely stable under educational expansion, with only a few exceptions (e.g., Sweden, the Netherlands, and Germany) where social origin differentials in educational attainment declined over decades (Breen et al 2009a, 2009b; Kerckhoff 1995a, 2001; Muller and Karle 1993; Pollak et al 2007; Shavit and Blossfeld 1993). As Shavit and Blossfeld (1993) concluded, “the rapid educational expansion did not reduce inequalities of educational opportunities”, and “the impact of educational reforms on changes in educational stratification seems to be negligible” (p. 21-22).

In places where the process of educational expansion differs from that of a typical industrial country, however, the general pattern of persistent or declined educational inequality does not apply. For instance, in Russia’s late- and post-Soviet periods and China’s early and recent reform periods, scholars find increased inequality at certain levels of education (Gerber 2000; Gerber and Hout 1995; Wu and Zhang 2010; Zhou et al 1998). In these settings, sudden changes in the distribution (and allocation) of educational opportunities took place in a short period of time.

The differences in the patterns of educational inequality (persistent or declined vs. increased) call for an integrated framework for theoretical explanations. Unfortunately, current theories in the field are mainly based on the rational choice approach (Breen and Goldthorpe 1997; Breen
and Yaish 2006; Breen et al 2009a, 2009b), and their explanations are mostly applicable to the inequality patterns found in industrial societies. The experience of educational inequality in industrializing societies is yet to be systematically examined and theorized.

Particularly, if patterns of educational inequalities are different in societies with different educational opportunity structures under educational expansion, we need to further investigate whether educational stratification in a specific country is conditional on its educational opportunity structure. Unlike the rational choice approach which usually ignores the structural and institutional features of educational expansion, the focus on the educational opportunity structure shifts our attention from individual educational choices to how structural and institutional conditions of the educational system affect educational attainment.

While most research in the literature focuses on the description of the trend of educational inequalities (Shavit and Blossfeld 1993, 2007), few studies have ever examined how the educational opportunity structure in a given society affects trends of educational inequality (Hout 2004). One reason is that it is uncommon to see a society where the educational opportunity structure differs from the typical pattern found in industrial countries along the course of educational expansion.

In this chapter, I fill in the gap by studying how the distribution and allocation of educational opportunities shape educational stratification in China’s reform period. As illustrated in Chapter 3, reforming China’s educational expansion has introduced both structural and institutional changes in the educational system that are rarely found elsewhere, making China a unique and valuable case for comparative educational research. Therefore, my aim of this chapter is not  

---

33In this dissertation, the terminology of “educational opportunity structure” refers to the distribution of educational opportunities in a given educational system. In Chapter 2, I have simulated various educational opportunities structures (Table 2.2, Chapter 2). My theoretical discussion and illustrative simulation suggest that the educational opportunity structure does affect patterns of educational inequalities.
solely to provide a description of the trends and patterns of educational stratification in reforming China, but also to supply a demonstration and test of the theoretical model in proposed in Chapter 2 and of the general hypothesis on educational attainment generated in Chapter 3.

Specifically, using the pooled national representative data of the 1996 Life History and Social Change survey and China General Social Survey (CGSS) 2005-2010, I study the different patterns of educational inequality in China during 1981-2010. What I find is that in the past three decades China has shown both patterns of increased and declined inequalities in schooling, and the various patterns are closely linked to the structural and institutional changes under educational expansion.

In the following sections, I will first briefly review the main literature in the field, and then formulate testable hypotheses based on the formal model in Chapter 2 and the general hypotheses in Chapter 3. Next, I will turn to the empirical analyses and demonstrate how social origins have affected children’s educational achievement under educational expansion. Then, I will test my theoretical explanations based on multilevel modelling results. Finally, I will discuss the implications of the Chinese patterns of educational inequality for comparative study of educational stratification as well as the trends of China’s labor market inequality.

LITERATURE REVIEW

A central topic in educational stratification research is whether inequality of educational opportunities changes along the course of educational expansion. With long-term growth in enrollments in school systems, it was once believed that people’s educational achievement would gradually become independent of their family backgrounds (Boudon 1974; Treiman 1970).

34 China’s economic reform started from 1978. However, the normalization of its educational system only began in early 1980s, particularly for college education (Zhou et al 1998). See also Footnote 26, Chapter 3.
However, as the expansion of education and the distribution of educational opportunities are two distinctive processes (Hauser and Featherman 1976; Mare 1980, 1981), empirical findings on trends in educational stratification based on cohort analysis had given little support to the claim (e.g., Featherman and Hauser 1978).

In most industrial societies, cross-country comparison suggests the general pattern is persistent inequality of educational opportunities since the early 20th century (Shavit and Blossfeld 1993): the effect of social origins on educational progress declined across transitions, and social origin differentials have basically remained unchanged across cohorts (e.g., Shavit and Blossfeld 1993; for recent reviews, see Buchmann and Hannum 2001; Breen and Jonsson 2005; Shavit and Blossfeld 2007). Even in some former state socialist countries where the state adopted an egalitarian policy that favored children from working classes (e.g., Hungary and Poland), the effects of family background on educational attainment were found to be stable (Hanley and McKeever 1997; Heyns and Biatecki 1993; Simkus and Andorka 1982; Szelenyi and Aschaffenburg 1993).

A pattern of declining educational inequality is also documented in countries such as the Netherlands (De Graaf and Ganzeboom 1993; De Graaf et al 2000), Germany (Muller and Karle 1993), Sweden (Erikson and Jonsson 1996; Jonsson et al 1996) and Norway (Lindbakk 1998). Recently, there has been growing evidences of declining educational inequality in Europe (Breen et al 2009a, 2009b). As Breen et al (2009, 2010) argue, a pattern of declining inequality may be more favorable than a pattern of persistent inequality, because “both primary and secondary effects change in ways such that declining disparities between classes in educational attainment can be expected” (2009b, p. 1480).
Considering the rising evidences of declining inequality, in a recent reassessment of the debate Shavit and Blossfeld (2007) contend that persistent inequality is probably wrong in its strong version, but a weaker version that allows for declining inequalities at lower educational levels and for declining inequalities in the middle of the 20th century can still stand. They point out that even among the results that report a reduction in educational inequality, the reduction is moderate. Moreover, if social origin is measured by parental education rather than father’s class, the relationship is still fairly constant over time. Therefore, they conclude that “persistent inequality still persists” by its weaker version (p. 52).

Yet, most findings of “persistent or declining inequality” are derived from countries where the educational system has gradually expanded over a long time. In such societies, usually, schooling cost was steadily reduced, parents’ subjective expectation on the benefits of education was slowly changed, and the increase in the enrollment started from lower levels and got to higher levels step by step (see Breen and Goldthorpe 1997 for related discussions). Under this type of educational system, according to rational choice theory (Breen and Goldthorpe 1997; Breen and Yaish 2006; Breen et al 2009a, 2009b), economic growth and educational expansion have reduced class differentials in children’s cognitive ability and educational aspiration, as well as in parents’ capacity to provide desirable education (usually through a uniform reduction of educational cost). In the meantime, different classes do not change their strategy of relative risk aversion with cohort replacement. Empirical results based on data from the Netherlands and West Germany supported these mechanisms (Becker 2003; Need and de Long 2001).

In contrast, in societies where educational expansion or contraction happens suddenly, educational inequalities at certain levels can be increased rather than persistent or declined. A sudden change in the educational policy could alter the allocation of educational resources
among different social groups. A quick-shift in the school system could also rapidly change children’s evaluation of educational costs and benefits, their educational preferences and aspirations, as well as the probabilities of schooling success. In this case, one may expect the patterns of educational inequality to be different from those of industrial societies.

Empirical evidences from transitional societies have lent some support for this point. In Soviet Russia, for example, when the rapidly-expanded secondary education created a “bottleneck” structure at the university level, the impact of social origins on access to university strengthened (Gerber and Hout 1995). Moreover, in the tumultuous late- and post-Soviet years, school enrollment was contracted from 1990 to 1994, and origin-based inequalities in access to academic secondary schools increased for cohorts who completed education in the period (Gerber 2000). Similarly, in China, educational contraction took place both at the college level during the Cultural Revolution (Deng and Treiman 1997) and at the senior high school level in the early reform periods (see also Figure 3.2, Chapter 3). Compared to the Maoist-era, results suggest that the effect of family class background on both secondary and tertiary educational transitions increased in the early reform periods (Zhou et al 1998). In addition, when educational expansion suddenly happened in the late 1990s, rising educational inequality at the college level is documented (Guo and Wu 2009; Wu and Zhang 2010).

The pattern of increased educational inequality is not well-predicted by the rational choice theory. It is also not fully accounted by other theoretical propositions which focus on the distribution and differentiation of educational opportunities. The “maximally maintained

---

35 Current studies on China’s educational inequality are mostly descriptive (Guo and Wu 2009; Wu and Zhang 2010). Some studies have even mistaken the basic trends (e.g., Li 2003, 2010). Even though recently there have been efforts to explain the Chinese pattern (Jiang and Tam 2015), their analytical strategy and results are not convincing. In addition, few studies have ever considered China’s particular educational opportunity structure, and how it contributes to understanding the general literature in comparative studies. To remedy these limitations, this chapter uses multiple sources of available data to present the basic patterns of China’s educational inequality in the past three decades and to provide a theoretical explanation based on the general theory in Chapter 2.
inequality” hypothesis claims that social origin differentials would only begin to decline when school enrollment at a given level has reached saturation (Raftery and Hout 1993). As the hypothesis assumes that increased enrollment would only preserve the previous origin-specific relationships unless a certain level of education becomes universal, it has a problem in predicting the rise of socioeconomic effects on school continuation. The “effectively maintained inequality” hypothesis further posits that even when quantitative enrollment at a given level is universal, advantaged groups can still look for qualitatively better education (Lucas 2001). Although it does articulate the possibility of rising educational inequality, it is more applicable to understanding inequalities in educational tracking. Moreover, it does not provide a mechanism to explain why the inequality could be “effectively maintained”.

Other than the existing theories in the field, the various empirical patterns suggest that educational opportunity structure can be a key factor that shapes the educational stratification process. In fact, in all the above cases where there is a pattern of “increased inequality”, the distribution of educational opportunities changed in a relatively short time, which was usually driven by state policy (Deng and Treiman 1997; Zhou et al 1996). Notably, either the contraction of school enrollments in Russia and China or the “bottleneck” structure in Soviet Russia is rarely seen in a typical industrial society.

In this Chapter, I present evidence from reforming China to understand how changes in the educational opportunity structure have implications on the patterns of educational inequality. As Chapter 3 shows, China has a special national enrollment planning system operated by the state, and its educational system is characterized by a “bottleneck” structure and rising educational costs under educational expansion. All these features are uncommon in a typical industrial society.
Moreover, China has a long historical tradition of using merit-based educational system (the Imperial Examination System) to select civic servants, which has made educational attainment not only a source of honor for the family, but also an effective avenue for social upward mobility (Chang 1955; Gan 2008). Due to this tradition, Chinese citizens are more likely than their counterparts in many other societies (including the US and Japan) to believe that merit-based education is one main routine for socioeconomic success (Whyte and Han 2008; Whyte and Guo 2009). Because of the significance of the perceived role of education for social upward mobility in China, it is also interesting to see how the changing educational opportunity structures affect the role of education for social mobility, especially for the disadvantaged groups.

**UNDERSTANDING EDUCATIONAL EXPANSION IN REFORMING CHINA**

In Chapter 2, by integrating both micro- and macro-based approaches, I have formally proposed a general model of educational expansion. With regard to the patterns of educational inequality, it provides at least three propositions. Frist, children from different social groups have different distributions of ability and resources (θ and φ parameters). Other things being equal, larger group differentials in these aspects will lead to more unequal distribution of educational opportunities among them (Table 2.2, Chapter 2).

Second, the structural and institutional changes introduced by educational expansion in the educational system have impacts on how the newly-available educational opportunities are allocated among different social groups (Δs and Δw parameters in the language of Chapter 2). Different allocation processes will result in different patterns of educational inequality.

Third, there is a linkage between educational inequalities at different levels under educational expansion. The relative sizes of the expansion rates (Ek and Ek+1 parameters) will
predict different patterns of educational inequality at the higher level (Table 2.2, Chapter 2).

Back to the case of reforming China, there are certain specific features of educational expansion. As summarized in Chapter 3, there was an increasing supply of junior high school graduates in the 1980s and 1990s, but the senior high school opportunity was declining in the 1980s and was almost stable in the 1990s. After the 1999 massive educational expansion, while the senior high school opportunity increased quickly but was still limited, the college opportunity increased dramatically (Figure 3.2, Chapter 3). As a result, China’s educational opportunity structure is featured by a “bottleneck” structure for senior high school attainment in the 1990s (similar to Soviet Russia, see Gerber and Hout 1995) and a more constricted “bottleneck” structure for college attainment after the 1999 expansion. In either case, the constraint was at the senior high school level: the transition rate to senior high school was about 22 percent in the 1990s and was 47.8 percent in 2010 (Table 3.1, Chapter 3).

Therefore, according to the third proposition, in the 1990s, the expansion rate of junior high school was larger than that of senior high school (i.e. the first scenario in Chapter 2), which would always predict a higher educational inequality at the senior high school level. Therefore, I expect that:

*Hypothesis 1: Educational inequality at the senior high school level increased in the 1990s.*

In addition, after the 1999 expansion, the expansion rate of senior high school was lower than that of college (i.e. the case of $E_k < E_{k+1}$ discussed in Chapter 2), suggesting a more constrained “bottleneck” structure after the 1999 expansion. In Chapter 3, I have extensively discussed the allocation of college opportunities under this educational opportunity structure. Combined with the rising educational cost for college, which means increasing social group differentials in resources (or a lower $\phi$ parameter in Chapter 2) and thus larger educational
inequality according to the first proposition, I expect that:

**Hypothesis 2: Educational inequality at the college level increased after the 1999 educational expansion.**

Therefore, my theoretical framework suggests a structural story (contracted educational opportunities) for increasing educational inequality at the senior high school level in the 1990s. It also suggests both a structure (“bottleneck” opportunity structure) and an institutional (rising college cost and larger resource differentials among social groups) story for rising inequality at the college level. The framework does not specifically consider how (expected) labor market returns to education affect the changing distribution of educational opportunities, as the labor market educational returns are not solely determined by educational expansion (see Footnote 6, Chapter 2).

Recently, Jiang and Tam (2015) have proposed a competition model to understand China’s rising educational inequality during 1988-2002. They argued that industrialization and educational expansion fail to account for the rising inequality in college attainment. Instead, their results lend support to a “market incentive hypothesis”, which argues that with a rising college earning premium, there are market incentives to intensify the competition for college and consequentially enlarge class differences in the admission success. However, this hypothesis is implausible to explain the effect of the 1999 massive educational expansion on educational inequality at the college level.

First, the last year of the time period of their study (1988-2002) is only three years after the 1999 educational expansion. Students who went to college during 1999-2002 have not yet entered the labor market in 2002. Thus the college income premium information in 1988-2002 held only for the pre-expansion cohorts, and the after-expansion cohorts of 1999-2002 had no
reliable information on their future college premium to justify their educational decisions.\endnote{36} Even if they were “myopia” decision makers, believing that college premium would become larger once they entered the labor market and then competing for it, the level of their competition would be offset by college expansion, as colleges became less selective after the expansion. Therefore, a competition model based on the “market competition hypothesis” cannot fully explain why the 1999-2002 cohorts still experienced increased educational inequality even if educational expansion lowered the level of competition among senior high school graduates.

Second, the trend of increasing college premium presented in their analyses is unlikely to hold after 2002. As my discussion in Chapter 3 suggests, if labor market demand does not change significantly, there is a tendency of declining labor market returns to education. Given the sharp increase of college graduates in a short period of time (Figure 3.6, Chapter 3), it would be misleading to assume that the labor market demand of college graduates exceeds the supply. In fact, as will be shown in Chapter 6 (Table 6.6), college premium declined by 28.7 percent during 2003-2008 in urban China.

Third, the measurement of educational expansion in Jiang and Tam (2015) is inappropriate. They use the “pool size”, defined as the ratio of upper secondary school graduates to the size of age cohort, to measure educational expansion, but this is not a measure of college expansion at all.\footnote{37} In the Chinese educational system, the yearly college enrollment is predetermined by the

\footnote{36}{It is possible that students who went to college during 1999-2002 used the college income premium in the admission year to expect their future college premium (Lucas 2001). It is also possible that students made the decision to go to college at the time of transition to senior high school (Breen and Yaish 2006; Breen et al 2009a), and expected their future college premium based on the labor market returns when aged 15 (Cameron and Heckman 1998; Jiang and Tam 2015). In either scenario, the 1999-2002 post-expansion students are “myopia” decision-makers. They are assumed to ignore the dramatic structural and institutional changes associated with the 1999 educational expansion. This is not a favorable assumption.}

\footnote{37}{The measure only captures the supply of senior high school graduates (including vocational school graduates) for college admission, or the size of the “competition pool” for college. It is not even a suitable measure for the level of educational expansion.}
national enrollment planning system ahead of the *College Entrance Exam* (Hannum et al. 2012; Shen 2011). Therefore, senior high school graduates’ opportunities for college do not depend on the size of the competition pool, but rather on how selective the state decides to make college admission each year.

In a word, Jiang and Tam (2015) have overstated the explanatory power of the “market incentive hypothesis” and underestimated the effect of educational expansion on college attainment. Given the vast change in China’s educational opportunity structure, it is premature to argue that educational expansion is irrelevant to patterns of changing educational inequality in reforming China. Therefore, to empirically test my interpretation of the rising inequality in college attainment after the expansion, I hypothesize that:

**Hypothesis 3a:** The larger the differences between college and senior high school transition rates, or the more constricted the bottleneck structure, the larger the effects of social origin on college attainment.

**Hypothesis 3b:** The higher the college cost, the larger the effects of social origin on college attainment.

Hypothesis 3a considers not only the enlarged college opportunity at the college level, but also how the allocation of college opportunity is shaped by the schooling opportunity provided for senior high students. It suggests that China’s “bottleneck” structure has significantly blocked disadvantaged social groups from college opportunities (see Figure 3.5, Chapter 3). Hypothesis 3b further contends that the rising college cost has enlarged role of the resource differentials among social groups, leading to a more unequal distribution of college opportunities.

_Hypotheses_ 3a and 3b do not rule out other possible explanations. Another explanatory competition, as China’s school admission is not analogous to a free market model. Given the dramatically increased transition rate since 1999 (Table 3.1, Chapter 3), the level of competition decreased even if the pool size increased.
factor could be the changing labor market returns to college (Jiang and Tam 2015). However, it is not crystal clear regarding the direction and mechanisms of the effect of labor market outcomes on college attainment. On the one hand, higher college premium could increase the competition for college and thus educational inequality (Jiang and Tam 2015); on the other hand, declining college premium could also exacerbate the social origin differentials in college opportunities, because the deflation of college credentials would post a more serious challenge to disadvantageous social groups, provided that they have already found it more difficult to pay for college. Therefore, the college premium should not be treated as a predictor but rather a product of two combined driving forces for educational inequality, i.e. market transition and educational expansion. In this sense, I do not provide explicit hypotheses regarding the effect of college premium on higher education inequalities.

In all three hypotheses, educational inequality is captured by social origin differentials in school attainment. In the Chinese context, one of the most important social origin factors is the household registration (hukou) origin (Chan and Zhang 1999; Knight and Song 1999; Whyte 2010a, 2010b; Wu and Treiman 2007; Zhang 2010). In the current educational system, primary, middle and high school admission is operated based on the admission district where the schools and students’ families are located (Hannum and Wang 2006). The majority of rural students can only go to nearby villages, towns and counties for primary and secondary education. Migrant children have the possibility of receiving primary and secondary education in cities, but their school enrollment status (xueji) is administered by the local educational authorities in the counties where their household registration belonged, and they have to take the College Entrance Exam back at their registered permanent residences. The geographic constraint of school admissions has systematically affected rural children’s educational opportunity structure. They
quality of their schools also differ significantly from that of their urban counterparts’. Therefore, household registration origin is one fundamental factor that shapes the structure of educational inequalities in the Chinese society (Hao et al 2014).

Another significant social origin factor is father’s class status. Although there are still debates on how to understand the class structure in contemporary China (e.g., Bian 2000; Li 2003; Zhou 2000a, 2000b), more and more evidence suggests that class has become increasingly important dimension to determine patterns of social stratification and inequality in Chinese society (Wu and Treiman 2007). For comparative purposes, in the dissertation, I adopt the Weberian EGP class schema (Erikson et al 1979, 1982, 2010) to characterize China’s class structure.38

DATA, VARIABLES AND METHODS

Data

The data for this chapter are drawn from the pooled samples of the 1996 Life History and Social Change Survey and four waves of the China General Social Survey (CGSS), 2005-2010. All the surveys employed a multi-stage stratified sampling design to obtain nationwide comparable random samples of adults (Treiman and Walder 1998). Since the fieldwork of the five surveys was mainly conducted by the same investigator, the Department of Sociology, Renmin University of China, the sampling schemes were fairly consistent (Bian and Li 2012;

38 It is still not clear whether the EGP class schema is the best way to characterize the structure of social class in reforming China (Lin 2008). Future work needs to be done to set up a solid foundation for the theoretical and empirical interpretation of China’s class structure.
In each of the surveys, counties in the chosen provinces were stratified by income levels and randomly selected based on a weighted sampling scheme. To balance the samples, the provincial capital and a lower-income city were also selected. Within each county and city, both urban and rural communities were all chosen randomly, and an adult from each sampled household was selected based on a KISH grid.

The 1996 *Life History and Social Change* data and the 2005-2010 CGSS data are particularly suitable for studying education-based social stratification and inequality in reforming China. First, they present a series of highly comparable cross-sectional observations of the dramatic social changes in the Chinese society. Particularly, since the first sample was collected in 1996 and the latest publicly available wave of the CGSS survey was implemented in 2010, eleven years after the 1999 educational policy, the series of data give researchers an opportunity to study the before and after effects of China’s massive educational expansion. Secondly, the five surveys gathered extensive information on respondents’ educational and occupational histories. In most of the surveys, respondents were asked to report detailed information of each level of education they have received (e.g., the start and end years, the enrollment status) as well as each job they have gone through (e.g., occupation, work unit). Therefore, recording the life trajectories of the respondents in terms of education and occupation was possible. Thirdly, information on family socioeconomic characteristics, including parents’ occupation, education and *hukou* status, was also available in various survey years. An overview of the major sources of data and the measurement of the available main variables used in this study is provided in the Appendix.

---

39 Communities are fixed in certain waves of the CGSS data. For example, in CGSS 2005 and 2006, the primary sampling units are the same. There is also a 2003 wave of the CGSS data, but that wave was limited to the urban population only. For more details for five datasets and their sampling procedures, see Treiman and Walder 1998, Wang 2011, and [http://www.chinagss.org/].
Using the pooled sample of the five datasets,\textsuperscript{40} I focus on those who have experienced their educational transitions during the reform periods, 1981-2010. I choose 1981 as the starting year because the normalization of the educational system only happened in the early 1980s (see Footnote 26, Chapter 3). Since in each survey there are certain youngest cohorts who are underrepresented by the data (usually few cases with year of transition close to the survey year), I exclude them based on the national statistics of the transition rates. This has yielded a sample of 13,534 and 6,025 cases for transitions to upper secondary school and tertiary school, respectively.

\textit{Variables}

Following the standard research design in the literature (e.g., Gerber and Hout 1995; Mare 1981), I define five levels of conditional transitions: from primary to junior high school, completion of junior high school, from junior high school to senior high school and/or vocational school, completion of senior high school, and from senior high school to college (including junior college, or \textit{dazhuan}). Given the fact that compulsory education has been universal since the mid-1990s, I focus on the determinants of post-compulsory transitions.

For each of the transitions, I derive the year of respondents’ educational transition from their report on educational history, or the enrollment status and specific grade if educational history information is unavailable. A caveat is that the year of college transition refers specifically to the departure year of senior high school, as some may choose to resume part-time college after

\textsuperscript{40} I use the pooled sample of the five datasets rather than one single dataset for this study. Compared to a single sample, the pooled sample provides not only more cases of those who have been affected by the 1999 educational expansion, but also a more accurate description of the actual social changes since the late 1990s. Moreover, for each of the topics studied in this dissertation, I have done sensitivity tests using different combinations of datasets whenever possible. Results based on one dataset or combined datasets do not change the basic empirical patterns obtained from the pooled sample. These results are available upon request.
starting to work (Li and Walder 2001). To indicate the effect of the 1999 massive educational expansion, I divide the years of post-compulsory transitions into three periods: 1981-1992, 1993-1998, and 1999-2010. The year 1992 witnessed Deng Xiaoping’s tour to southern China, after which China deepened its economic reforms and marketization. The year 1999 marked the launch of an educational policy that led to quick expansion in senior high school and dramatic expansion in higher education. 41

The main individual-level independent variables are social origin measures, as indicated by father’s education, father’s occupation, and hukou origin. Father’s education is measured by years of schooling and educational levels (1=primary or below, 2=junior high school, 3=upper secondary school, 4=college). Father’s occupation is measured both by a continuous ISEI score (Ganzeboom et al 1992) and by a categorical measure of class status based on the EGP class schema (Erikson, Goldthorpe, and Potocarero 1979; Ganzeboom, Liujkx, and Treiman 1992; Ganzeboom and Treiman 1996, 2003). Specifically, I code father’s occupation into a 6-category version of the EGP scheme, which was previously applied to China (Wu and Treiman 2007). The relationship between the 10-category version proposed by Erikson et al (1979) and the 6-category version used in the dissertation is presented in Table 4.1.

Hukou origin is a dummy variable, indicating whether the respondent’s father’s hukou was urban when aged 14 or 18 (1=urban, 0=rural). If that information is not available such as in CGSS 2008, then I use respondents’ hukou change status to get the hukou origin measure. Because of the fundamental urban-rural divide in the Chinese societies (Wu and Treiman 2004, 2007), hukou origin captures not only one important aspect of social origins, but also one source of structural inequality in the educational system.

41 I have checked the robustness of the division of periods by using alternative cut-offs, which does not change the substantive findings. The general pattern is also robust when using birth cohorts to approximate periods.
Table 4.1: The Six-category EGP Schema

<table>
<thead>
<tr>
<th>Original classification</th>
<th>New classification</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. Large proprietors, higher professionals and managers</td>
<td>6</td>
</tr>
<tr>
<td>II. Lower professionals and managers</td>
<td>6</td>
</tr>
<tr>
<td>III. Routine nonmanual workers</td>
<td>5</td>
</tr>
<tr>
<td>IVa. Small proprietors with employees</td>
<td>4</td>
</tr>
<tr>
<td>IVb. Small proprietors without employees</td>
<td>4</td>
</tr>
<tr>
<td>V. Lower grade technicians and manual supervisors</td>
<td>3</td>
</tr>
<tr>
<td>VI. Skilled manual workers</td>
<td>3</td>
</tr>
<tr>
<td>VIIa. Unskilled and semiskilled manual workers</td>
<td>2</td>
</tr>
<tr>
<td>IVc: Self-employed farmers</td>
<td>1</td>
</tr>
<tr>
<td>VIIb. (Unskilled) agricultural workers</td>
<td>1</td>
</tr>
</tbody>
</table>

Source: Wu and Treiman 2007

Another important variable in the literature is the number of siblings (e.g. Cameron and Heckman 1998; Mare and Chen 1986; Sieben 2001), as a family’s available resources need to be distributed among all the children. Unfortunately, the information is biased in the CGSS 2005 and 2006 waves. Therefore, I only treat it as a control variable in the empirical analysis. Finally, gender, a dummy variable (male=1), is included in the models as another control variable.

Considering the possibility that the effect of hukou origin might differ by gender in the Chinese setting (e.g., Hannum and Xie 1994; Hannum 2005; Hannum et al 2012; Lavely et al 1990; Tong 2009), an interaction term between hukou origin and gender is included in the models.

To provide a direct assessment of Hypothesis 3a and 3b, I also include two year-specific contextual variables into the model building. The “bottleneck” structure is measured by the year-specific difference in the national transition rates to college and to senior high school. The rising college cost is indicated by the average tuition and fees for college students each year from 1995-2010, the only time points for which the national statistics are available. Both measures have been presented in Chapter 3 (Table 3.1 and Table 3.3).
Table 4.2 summarizes the descriptive statistics for individual-level independent variables, varied by post-compulsory school transitions. The table shows that the composition of family backgrounds for those who were at the risk of making transitions differs not only across periods, but also across transitions. Within each level of transition, the means and standard deviations of father’s years of schooling both increased across periods, suggesting that children who were in the risk set of making educational transitions have become more heterogeneous in terms of father’s education. The heterogeneity of father’s occupation also increased among potential college transition takers, as indicated by the standard deviations of father’s ISEI score. In
addition, the percentage of rural hukou holders among children who faced college transition decreased in 1999-2010. Therefore, the supply of senior high school graduates (for potential college transition) became more heterogeneous in father’s education and occupation after the 1999 expansion, with a larger share of urban hukou holders. In other words, children were socially included and excluded for educational opportunities (Arum et al 2007; Muller and Karle 1993).

**Methods**

I estimate sequential educational transition models to examine the impacts of family backgrounds on educational attainment (Hauser and Andrew 2006; Mare 1980, 1981, 2006, 2011). Specifically, these models take the following logit form:

\[
\log_e \left( \frac{p_{ij}}{1-p_{ij}} \right) = \beta_{j0} + \sum_k \beta_{jk} x_{ijk} \tag{1}
\]

where \(p_{ij}\) is the probability that the \(i^{th}\) individual will make the \(j^{th}\) transition, \(x_{ijk}\) is the value for the \(i^{th}\) individual deciding whether to make the \(j^{th}\) transition on the \(k^{th}\) independent variable, \(\beta_{j0}\) is the constant, and \(\beta_{jk}\) denotes the effect of a unit change in \(x_{ijk}\) on the log-odds of grade progression. Since the logit coefficient \(\beta_{jk}\) is hard to interpret (Breen et al 2011), I report the marginal effect instead, which indicates the rate of change in the probability \(p_{ij}\) with respect to one unit increase in \(x_{ijk}\).

The sequential educational transition models were once criticized for their implicit assumption of “myopia”, which assumes educational decisions are only based on resources available at one time point (Cameron and Heckman 1998). The inclusion of time-varying covariates in the models can partly solve the problem (Lucas 2001), even if myopia is an
unfavorable assumption. Since the pooled data analyzed in this chapter are based on multiple comparable cross-sectional surveys, which contain a time-varying component, the results do not suffer much from this methodological criticism.

Another criticism of school transition models is the assumption of unilinear sequential transitions, which cannot detect the multiple branches at certain transition levels that might have path dependence and tracking effects (Gamoran and Mare 1989; Meer 2007). To remedy this problem, Breen and Jonsson (2000) proposed a multinomial model to study educational transitions (see also Hillmert and Jacob 2003). Considering educational tracking in China (Figure 3.1, Chapter 3), I also use the multinomial model to study the social differentials in the transition to the academic and vocational tracks at the upper secondary level.

To examine the temporal trends, I estimate models in three historical periods divided by the years of educational transitions. To further see whether the effects of independent variables change across periods, a statistical test on comparing the coefficients of different models is implemented through z-score statistics (Clogg et al 1995), which is defined by:

\[
Z = \frac{\beta_1 - \beta_2}{\sqrt{SE\beta_1^2 + SE\beta_2^2}}
\]

In the formula, \(\beta_1\) and \(\beta_2\) are the coefficients rather than the marginal effects, and \(SE\beta_1\) and \(SE\beta_1\) are the standard errors of the coefficients.

To further test Hypothesis 3a and 3b, I estimate two-level hierarchical logistic regression models for the transition to college. The model is specified as follows:

\[
\text{Level-1 Model: } \log_e \left( \frac{p_{ij}}{1-p_{ij}} \right) = \beta_{0j} + \sum_k \beta_{kj} x_{kij} + \mu_{ij}
\]

According to Cameron and Heckman (1998), in the absence of time-varying covariates, sequential school transition models cannot be parametrically identified and the coefficients may be affected by the distribution of unobserved heterogeneity.
Level-2 Model: \[ \beta_{0j} = \gamma_{00} + \gamma_{01} z_j + \delta_{0j} \]
\[ \beta_{kj} = \gamma_{k0} + \gamma_{k1} z_j + \delta_{kj} \]

In the level-1 model, \( \beta_{0j} \) refers to the intercept terms for the \( j^{th} \) year of transition, \( \beta_{kj} \) refers to the coefficient for the \( i^{th} \) individual in the \( j^{th} \) year with an observed value on the \( k^{th} \) variable, and \( \mu_{ij} \) is the individual-level residual term. The level-2 model further estimates the coefficient \( \beta_{kj} \) in the level-1 model as a function of an additional contextual predictor in the \( j^{th} \) year (\( z_j \)) with an error term \( \delta_{kj} \). Therefore, the model is a random coefficient logit model, or a slopes-as-outcomes model (Raudenbush and Bryk 1986).

**RESULTS**

*Descriptive Statistics*

Table 4.3 presents a summary of the transition rates for different social groups in the three periods. As it indicates, despite the overall increase of the transition rates at all the educational levels (the first row), the opportunities of school expansion were not evenly distributed. Because of the successful implementation of the nine-year compulsory education, the between-group gaps in the rate of transition to junior high schools decreased in second period, and remained almost stable in the third period. For instance, the transition rate for rural *hukou* holders increased from 77.8 percent in 1981-1992 to 89.9 percent in 1993-1998, and then slightly to 91.3 percent in 1999-2010, whereas the transition rate for urban children was already high, increased only from 97.2 percent to 99.4 percent. The evidence suggests that educational expansion in compulsory education in the 1980s and the 1990s certainly helped children from rural/disadvantaged backgrounds, particularly in the first period.
Table 4.3: Transition Rates by Educational Transitions and Periods, the Pooled Sample, 1981-2010

<table>
<thead>
<tr>
<th>Transition Rates %</th>
<th>Primary to Junior High</th>
<th>Junior High to Senior High</th>
<th>Senior High to College</th>
<th>From Primary to College</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall</td>
<td>87.9 84.0 93.3 94.6</td>
<td>38.2 33.8 37.3 49.2</td>
<td>44.6 25.0 43.9 75.2</td>
<td>0.07 0.15 0.35</td>
</tr>
<tr>
<td>Hokou origin</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urban</td>
<td>97.9 97.2 98.8 99.4</td>
<td>57.8 51.5 58.6 68.6</td>
<td>53.1 31.0 51.4 83.8</td>
<td>0.16 0.30 0.57</td>
</tr>
<tr>
<td>Rural</td>
<td>82.5 77.8 89.9 91.3</td>
<td>24.7 22.7 23.3 31.9</td>
<td>30.2 16.4 30.6 50.6</td>
<td>0.03 0.06 0.15</td>
</tr>
<tr>
<td>Gap: Urban-Rural</td>
<td>15.4 19.4 8.9 8.1</td>
<td>33.1 28.8 35.3 36.7</td>
<td>22.9 14.6 20.8 33.2</td>
<td>0.13 0.23 0.42</td>
</tr>
<tr>
<td>Father’s education</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary or below</td>
<td>81.4 78.7 88.1 88.9</td>
<td>26.0 25.1 24.9 31.9</td>
<td>28.9 18.4 35.0 48.6</td>
<td>0.04 0.08 0.15</td>
</tr>
<tr>
<td>Junior high school</td>
<td>93.6 90.9 95.4 97.9</td>
<td>42.1 39.7 41.1 46.5</td>
<td>45.3 22.0 41.6 73.4</td>
<td>0.08 0.16 0.33</td>
</tr>
<tr>
<td>Senior high school</td>
<td>97.3 95.2 99.1 99.5</td>
<td>56.3 51.1 53.2 63.6</td>
<td>59.9 37.0 50.5 81.1</td>
<td>0.18 0.27 0.51</td>
</tr>
<tr>
<td>College</td>
<td>99.2 98.9 99.2 100.0</td>
<td>77.4 73.0 76.7 85.4</td>
<td>68.0 47.1 71.5 91.5</td>
<td>0.34 0.54 0.78</td>
</tr>
<tr>
<td>Gap: College-Primary</td>
<td>17.8 20.2 11.1 11.1</td>
<td>51.4 47.9 51.8 53.3</td>
<td>39.1 28.7 36.5 42.9</td>
<td>0.30 0.47 0.64</td>
</tr>
</tbody>
</table>

Note: Data are weighted. The gap is the difference in the transition rate between two groups.

a The probabilities are the products of respective rates in the first three panels.
The social group differentials in the transition to senior high school did not follow the same pattern. Clearly, children with disadvantaged backgrounds did not gain much of an opportunity to attend senior high school in the second period (1993-1998): the transition rate was almost the same as that in the first period for children with rural hukou origin. In contrast, it increased for urban children and for those whose father was highly educated or occupied a professional/managerial position. The urban-rural gap in the transition rate continued to increase in the third period, so do the gaps among other social origin groups characterized by father’s education and class. Moreover, it seems that children with median-level family backgrounds have benefited substantively in obtaining senior high school opportunities in the third period.

The transition to college follows a more pronounced pattern. Consistent with the national statistics, tertiary education expanded much faster than senior high school education. The overall rates of transition to college in the second and third periods surpassed the rates of senior high school transition. Yet, after the 1999 expansion, the distribution of the college opportunity has become more unequal. Although the transition rate became larger both for the urban and rural children in the third period, the gap enlarged: in 1993-1998, the urban-rural difference in the transition rate was 20.8 percent; in 1999-2010, it became 33.2 percent. The trend was similar for other social origin group differentials captured by father’s education and occupation. As a result, although college expansion increased the probability of children with disadvantaged backgrounds to enter college, it has benefited children from advantageous social groups more in the third period.
The last panel further calculates the cumulative probability of going to college starting from primary school. The results show that the rate of going to college in the third period was 5 times of that in the first period (.35/.07, the first row). Compared to rural children, urban children’s college opportunities grew much faster: their calculated probability of going to college conditional on finishing primary school increased by 3.5 times, from 0.16 to 0.57. In contrast, the probability for rural children only increased to 0.15 in the third period, which was still lower than that for urban children in the first period (0.16). Children with advantaged educational and class backgrounds also enjoyed the advantage in the cumulative likelihood of college transition in the third period. Given the universal availability of junior high school education, the low probabilities of college education for children from disadvantaged groups suggest that a substantial proportion of them were selected-out at the senior high school level, which reflects the “bottleneck” structure in China’s educational system.

**Changing Educational Inequalities**

Table 4.4 presents estimated marginal effects of the binary logit models on the sequential school transitions. Results show that family background variables affect not only the attainment of schools, but also the likelihood of finishing them. Other things being equal, children with urban *hukou* origin are significantly more likely to finish both junior and senior high schools than their rural counterparts. Having a father with better education or a professional/managerial position also helps the children to avoid dropping out from junior high schools, but only children with a college-educated father are more likely to complete senior high schools than those whose father has primary or no education.

---

43 The calculation does not consider the dropout rates of junior and senior high schools. According to the pooled sample, the average dropout rate for junior high was 12.9 percent in 1996, and around 8 percent in 2005-2010. The senior high school dropout rate was lower, with an average rate of 4.18 percent in the five datasets.
Table 4.4: Estimated Marginal Effects for Binary Logit Models on Sequential School Transitions, 1981-2010

<table>
<thead>
<tr>
<th>Independent Variables</th>
<th>Primary to Junior High</th>
<th>Complete Junior High</th>
<th>Junior to Senior High</th>
<th>Complete Senior High</th>
<th>Senior to College</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Hukou origin (urban=1)</td>
<td>0.085**</td>
<td>0.079**</td>
<td>0.351**</td>
<td>0.027*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.013)</td>
<td>(0.010)</td>
<td>(0.021)</td>
<td>(0.011)</td>
</tr>
<tr>
<td></td>
<td>Father’s education a</td>
<td>0.040**</td>
<td>0.019**</td>
<td>0.131**</td>
<td>-0.002</td>
</tr>
<tr>
<td></td>
<td>Junior high school</td>
<td>(0.006)</td>
<td>(0.005)</td>
<td>(0.018)</td>
<td>(0.004)</td>
</tr>
<tr>
<td></td>
<td>Senior high school</td>
<td>0.053**</td>
<td>0.022**</td>
<td>0.232**</td>
<td>0.008</td>
</tr>
<tr>
<td></td>
<td>(0.008)</td>
<td>(0.008)</td>
<td>(0.020)</td>
<td>(0.007)</td>
<td>(0.025)</td>
</tr>
<tr>
<td></td>
<td>College</td>
<td>0.054**</td>
<td>0.052**</td>
<td>0.443**</td>
<td>0.022**</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.012)</td>
<td>(0.005)</td>
<td>(0.033)</td>
<td>(0.005)</td>
</tr>
<tr>
<td></td>
<td>Father’s EGP class b</td>
<td>0.047**</td>
<td>0.024*</td>
<td>0.129**</td>
<td>0.010</td>
</tr>
<tr>
<td></td>
<td>Profs., managers (I,II)</td>
<td>(0.009)</td>
<td>(0.010)</td>
<td>(0.026)</td>
<td>(0.010)</td>
</tr>
<tr>
<td></td>
<td>Routine nonmanual (III)</td>
<td>0.040**</td>
<td>-0.004</td>
<td>0.093**</td>
<td>0.010</td>
</tr>
<tr>
<td></td>
<td>(0.009)</td>
<td>(0.013)</td>
<td>(0.030)</td>
<td>(0.006)</td>
<td>(0.043)</td>
</tr>
<tr>
<td></td>
<td>Small owners (IVa,IVb)</td>
<td>0.021</td>
<td>0.006</td>
<td>0.064*</td>
<td>-0.015</td>
</tr>
<tr>
<td></td>
<td>Foremen, skilled (V,VI)</td>
<td>0.029**</td>
<td>0.014</td>
<td>0.016</td>
<td>-0.001</td>
</tr>
<tr>
<td></td>
<td>(0.006)</td>
<td>(0.009)</td>
<td>(0.030)</td>
<td>(0.008)</td>
<td>(0.032)</td>
</tr>
<tr>
<td></td>
<td>Semi- &amp; unskilled (VIa)</td>
<td>0.027**</td>
<td>0.005</td>
<td>0.014</td>
<td>-0.005</td>
</tr>
<tr>
<td></td>
<td>(0.009)</td>
<td>(0.007)</td>
<td>(0.025)</td>
<td>(0.011)</td>
<td>(0.035)</td>
</tr>
<tr>
<td></td>
<td>Number of siblings</td>
<td>-0.008**</td>
<td>-0.001</td>
<td>-0.023**</td>
<td>-0.003*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.006)</td>
<td>(0.001)</td>
</tr>
<tr>
<td></td>
<td>Sex (male=1)</td>
<td>0.020**</td>
<td>-0.001</td>
<td>0.052**</td>
<td>-0.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.006)</td>
<td>(0.006)</td>
<td>(0.016)</td>
<td>(0.006)</td>
</tr>
<tr>
<td></td>
<td>Sex*Hukou origin</td>
<td>0.001</td>
<td>-0.012</td>
<td>-0.013</td>
<td>0.013</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.017)</td>
<td>(0.014)</td>
<td>(0.025)</td>
<td>(0.011)</td>
</tr>
<tr>
<td></td>
<td>Period c</td>
<td>0.040**</td>
<td>0.012*</td>
<td>0.055**</td>
<td>0.008</td>
</tr>
<tr>
<td></td>
<td>1993-1998</td>
<td>(0.004)</td>
<td>(0.006)</td>
<td>(0.018)</td>
<td>(0.006)</td>
</tr>
<tr>
<td></td>
<td>1999-2010</td>
<td>0.041**</td>
<td>0.011</td>
<td>0.149**</td>
<td>0.010</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.011)</td>
<td>(0.008)</td>
<td>(0.022)</td>
<td>(0.008)</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>12,917</td>
<td>11,670</td>
<td>10,998</td>
<td>6,090</td>
</tr>
<tr>
<td></td>
<td>Log-likelihood</td>
<td>-4079.1</td>
<td>-2911.4</td>
<td>-5852.0</td>
<td>-811.46</td>
</tr>
<tr>
<td></td>
<td>Pseudo R-square</td>
<td>0.155</td>
<td>0.092</td>
<td>0.198</td>
<td>0.062</td>
</tr>
</tbody>
</table>

Note: Robust standard errors in parentheses. Data are weighted. Within-province clustering effects are controlled.  
* p<.05; ** p<.01 (two-tailed test)  
a Refer to “primary school or below”.  
b Refer to “farmers (IVc, VIIb)”.  
c Refer to “1981-1992”
In addition, comparing the models across transitions, we see that the effects of family background are the highest for the transition to senior high school. Other things being equal, rural children were disadvantaged by 35.1 percent in the probability of transition to senior high school over urban children, whereas their probabilities of transition to junior high school and to college were 8.5 percent and 23.8 percent less. Similarly, the effects of father’s education are the most salient in the transition to senior high school. With regard to the interaction term between sex and *hukou* origin, it is only significant for the transition to college. Its negative effect suggests that urban males’ advantage in going to college over urban females’ is significantly lower than that of rural males over rural females’.

The pattern that family background effects are the highest at the upper secondary school level is rarely anticipated by the existing literature. Rather, the general pattern in comparative studies is a pattern of waning effects of family background on school transitions (see Shavit and Blossfeld 1993: 6-10 for a summary), which is explained based on the life course perspective or the “maximally maintained inequality” hypothesis (Breen and Jonsson 2005: 236; Lucas 2001). In the Chinese case, I suppose an explanation based on the educational opportunity structure is more plausible, as senior high school education has become the “bottleneck” for school transitions.

In Table 4.5, I further examine educational inequalities in the transitions to senior high school and to college, separated by the three periods in which the transitions occurred. To test whether the effects significantly differ across periods, z-score statistics (Paternoster et al 1998) based on the estimated coefficients and standard errors for each model are calculated.
Table 4.5: Estimated Marginal Effects for Binary Logit Models on Post-compulsory Transitions, by Periods

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Model 1</td>
<td>Model 2</td>
<td>Model 5</td>
<td>2 vs. 1 3 vs. 2</td>
<td>Model 1</td>
<td>Model 2</td>
<td>Model 5</td>
<td>2 vs. 1 3 vs. 2</td>
</tr>
<tr>
<td>Hukou origin (urban=1)</td>
<td>0.291**</td>
<td>0.407**</td>
<td>0.358**</td>
<td>2.96** -1.27</td>
<td>0.161**</td>
<td>0.153**</td>
<td>0.298**</td>
<td>-0.80 2.60**</td>
</tr>
<tr>
<td>(0.027)</td>
<td>(0.036)</td>
<td>(0.036)</td>
<td></td>
<td></td>
<td>(0.038)</td>
<td>(0.050)</td>
<td>(0.049)</td>
<td></td>
</tr>
<tr>
<td>Father’s education *</td>
<td>0.126**</td>
<td>0.133**</td>
<td>0.117**</td>
<td>0.12 -0.05</td>
<td>0.032</td>
<td>-0.009</td>
<td>0.070</td>
<td>-0.83 1.67</td>
</tr>
<tr>
<td>(0.021)</td>
<td>(0.023)</td>
<td>(0.036)</td>
<td></td>
<td></td>
<td>(0.025)</td>
<td>(0.038)</td>
<td>(0.035)</td>
<td></td>
</tr>
<tr>
<td>Junior high school</td>
<td>0.239**</td>
<td>0.248**</td>
<td>0.172**</td>
<td>-0.01 -0.59</td>
<td>0.151**</td>
<td>0.071</td>
<td>0.108**</td>
<td>-1.64 1.13</td>
</tr>
<tr>
<td>(0.030)</td>
<td>(0.037)</td>
<td>(0.034)</td>
<td></td>
<td></td>
<td>(0.033)</td>
<td>(0.045)</td>
<td>(0.031)</td>
<td></td>
</tr>
<tr>
<td>Senior high school</td>
<td>0.391**</td>
<td>0.395**</td>
<td>0.260**</td>
<td>0.18 -1.02</td>
<td>0.209**</td>
<td>0.260**</td>
<td>0.268**</td>
<td>0.32 0.40</td>
</tr>
<tr>
<td>(0.031)</td>
<td>(0.039)</td>
<td>(0.087)</td>
<td></td>
<td></td>
<td>(0.068)</td>
<td>(0.069)</td>
<td>(0.040)</td>
<td></td>
</tr>
<tr>
<td>College</td>
<td>0.113**</td>
<td>0.134**</td>
<td>0.192**</td>
<td>0.56 1.89</td>
<td>0.026</td>
<td>0.072</td>
<td>0.182*</td>
<td>0.51 2.12*</td>
</tr>
<tr>
<td>(0.031)</td>
<td>(0.046)</td>
<td>(0.048)</td>
<td></td>
<td></td>
<td>(0.035)</td>
<td>(0.049)</td>
<td>(0.039)</td>
<td></td>
</tr>
<tr>
<td>Routine nonmanual (III)</td>
<td>0.079*</td>
<td>0.139**</td>
<td>0.155*</td>
<td>1.06 0.98</td>
<td>-0.026</td>
<td>0.042</td>
<td>0.044</td>
<td>1.11 0.32</td>
</tr>
<tr>
<td>(0.038)</td>
<td>(0.056)</td>
<td>(0.068)</td>
<td></td>
<td></td>
<td>(0.053)</td>
<td>(0.063)</td>
<td>(0.045)</td>
<td></td>
</tr>
<tr>
<td>Small owners (IVA,IVb)</td>
<td>0.015</td>
<td>0.083</td>
<td>0.110**</td>
<td>1.81 0.44</td>
<td>-0.058</td>
<td>0.044</td>
<td>0.051</td>
<td>1.47 0.15</td>
</tr>
<tr>
<td>(0.040)</td>
<td>(0.049)</td>
<td>(0.038)</td>
<td></td>
<td></td>
<td>(0.042)</td>
<td>(0.091)</td>
<td>(0.047)</td>
<td></td>
</tr>
<tr>
<td>Foremen, skilled (V,VI)</td>
<td>-0.038</td>
<td>0.104**</td>
<td>0.132**</td>
<td>1.65 0.31</td>
<td>-0.018</td>
<td>0.018</td>
<td>0.015</td>
<td>0.57 -0.34</td>
</tr>
<tr>
<td>(0.043)</td>
<td>(0.045)</td>
<td>(0.044)</td>
<td></td>
<td></td>
<td>(0.036)</td>
<td>(0.074)</td>
<td>(0.030)</td>
<td></td>
</tr>
<tr>
<td>Semi- &amp; unskilled (VIIa)</td>
<td>-0.012</td>
<td>0.033</td>
<td>0.058</td>
<td>0.97 0.68</td>
<td>-0.065</td>
<td>-0.015</td>
<td>-0.048</td>
<td>0.78 -0.20</td>
</tr>
<tr>
<td>(0.032)</td>
<td>(0.045)</td>
<td>(0.040)</td>
<td></td>
<td></td>
<td>(0.041)</td>
<td>(0.055)</td>
<td>(0.044)</td>
<td></td>
</tr>
<tr>
<td>Number of siblings</td>
<td>-0.019**</td>
<td>-0.029*</td>
<td>-0.014</td>
<td>-0.61 0.63</td>
<td>-0.009</td>
<td>-0.027</td>
<td>0.006</td>
<td>-0.73 1.80</td>
</tr>
<tr>
<td>(0.007)</td>
<td>(0.013)</td>
<td>(0.015)</td>
<td></td>
<td></td>
<td>(0.007)</td>
<td>(0.015)</td>
<td>(0.010)</td>
<td></td>
</tr>
<tr>
<td>Sex (male=1)</td>
<td>0.056*</td>
<td>0.085*</td>
<td>0.000</td>
<td>0.41 -1.03</td>
<td>0.077**</td>
<td>0.122*</td>
<td>0.070*</td>
<td>0.39 -0.41</td>
</tr>
<tr>
<td>(0.023)</td>
<td>(0.039)</td>
<td>(0.040)</td>
<td></td>
<td></td>
<td>(0.027)</td>
<td>(0.053)</td>
<td>(0.031)</td>
<td></td>
</tr>
<tr>
<td>Sex*Hukou origin</td>
<td>0.033</td>
<td>-0.122*</td>
<td>0.016</td>
<td>-2.72** 1.67</td>
<td>-0.059</td>
<td>-0.041</td>
<td>-0.182**</td>
<td>0.51 -2.42**</td>
</tr>
<tr>
<td>(0.028)</td>
<td>(0.051)</td>
<td>(0.072)</td>
<td></td>
<td></td>
<td>(0.035)</td>
<td>(0.065)</td>
<td>(0.040)</td>
<td></td>
</tr>
</tbody>
</table>

N | 6,074 | 2,510 | 2,414 | 2,606 1,261 1,762 |
Log-likelihood | -928.32 | -422.08 | -369.34 | -1114.45 -675.3 -714.9 |
Pseudo R-square | 0.125 | 0.223 | 0.127 | 0.064 0.058 0.105 |

Note: Robust standard errors in parentheses. Data are weighted. Within-province clustering effects are controlled. * p<.05; ** p<.01 (two-tailed test) a Refer to “primary school or below”. b Refer to “farmers (IVc,VIb)”. c Reported are Z-statistics testing coefficient differences across models.
As we can see from the table, the marginal effects of hukou origin and father’s education on the transition to senior high school follow an inverted-U-shaped across periods (first increased and then decreased), but the effect of father’s class increased in both the second and the third periods. Specifically, controlling for other variables, the probability of transition to senior high school for rural children was 29.1 percent, 40.7 percent, and 35.8 percent less than for urban children in the three periods, respectively, with the size of the effect highest in the second period. Similarly, net of other factors, children whose father has a college credential are 39.1 percent more likely to go to senior high school in the first period than those whose father has primary or no education. The figure increased to 39.5 percent in the second period, but decreased to 26.0 percent in the third period. The effect of father’s class also becomes larger in the second period. Other things being equal, children whose father was a professional or manager were 13.4 percent higher in the probability of transition to senior high school than children with a farm origin in the second period; in the first and third periods, their advantages are 11.3 percent and 19.2 percent higher.

The effect difference test further shows that the estimated marginal effects of hukou origin in the first two periods significantly differ from each other (z=2.96), but the effect difference between those in the second and the third periods is not statistically significant (z=-1.27). In addition, changes in the effects of father’s education and class are not statistically significant across periods, but the direction is that social origin effects all increased in the second period. Therefore, Hypothesis 1 which expects increasing educational inequality at the senior high school level in the 1990s is supported by the data.

Finally, the effect of the interaction term between sex and hukou origin is significant only in the second period, showing that in that period while rural males enjoyed a 8.5 percent advantage
over rural females, urban males experienced a 3.7 (0.085-0.122) percent disadvantage over urban females. The effect difference is significant for the interaction term in the second period ($z=-2.72$). The result confirms previous observations of China (Bauer et al 1992; Hannum 2002, 2005; Hannum and Xie 1994; Lavelle et al 1990) and is consistent with findings elsewhere: “males’ advantage over females in education has gradually disappeared (Buchmann et al 2008); in some cases, it has reversed” (Gerber and Hout 1995: 612).

With respect to the transition to college, it is clear that the effects of hukou origin, father’s education and class position are all greater in the third period (1999-2010). Compared to the first two periods, rural children became more disadvantaged in the transition to college: other things being controlled, their probability of entering college was 29.8 percent less than for urban children, and the effect difference was significant ($z=2.60$). The effect of father’s education on college transition also enlarges, although the effect is not statistically significant ($z=0.40$). Compared to children whose father has primary or no education, having a college-educated father predicts an average increase in the probability of going to college from 26.0 percent higher in the second period to 26.8 percent higher in the third period. At the same time, class background effects became more prominent in the third period. In particular, net of other factors, the probability of transition to college for children of professionals/managers was 18.2 percent significantly higher than for farmers’ children, and the effect difference was also significant in the last two periods ($z=2.12$).

These results indicate that social origin effects on college attainment have been significantly increased after the 1999 expansion. Particularly, urban hukou origin and father being a manager/professional predict significantly higher probabilities of going to college after the expansion. Father’s education also has a positive effect on college attainment, although the effect
is not significantly larger in the third period. These results provide strong support for Hypothesis 2.

In addition, with regard to the effect of gender, it is positively significant in the first two periods. In the third period, the interaction effect between sex and hukou origin also becomes significant, with a z-score of -2.42. Therefore, while rural males are 7.0 percent more likely to obtain college education than rural females after the 1999 expansion, their urban counterparts are 11.2 percent (0.07-0.182) less likely to go to college than urban females. Again, males’ advantage for school attainment has been reversed for urban children, but not for rural children.

To capture the academic vs. vocational tracking system at the upper secondary level, Table 4.6 summarizes the estimated marginal effects of multinomial logit models on upper secondary educational transitions. Similar to the results in Table 4.5, rural children had the largest disadvantages (27.1 percent lower in the probability) over urban children in access to the academic track of upper secondary schools in the second period (1993-1998). The effect difference test reveals that the hukou origin effects differed significantly in the first two periods (z=3.83); in the second and third periods, the effect difference was insignificant (z=-1.24). In a similar vein, the marginal effects of father’s education on academic track transition were the largest in the second period, but differences of the effects across periods are nonsignificant. Having a father with a managerial/professional position predicts increasingly larger probability of going to the academic track in the second and third periods, but the changes in the effect across periods are statistically nonsignificant.
Table 4.6: Estimated Marginal Effects for Multinomial Logistic Regression on Upper Secondary School Transitions, by Periods

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
</tr>
<tr>
<td><strong>Hukou origin (urban=1)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Academic Vocational</td>
<td>0.202**</td>
<td>0.079**</td>
<td>0.271**</td>
<td>0.078**</td>
</tr>
<tr>
<td></td>
<td>(0.022)</td>
<td>(0.016)</td>
<td>(0.028)</td>
<td>(0.029)</td>
</tr>
<tr>
<td>Academic Vocational</td>
<td>0.271**</td>
<td>0.078**</td>
<td>0.262**</td>
<td>0.013</td>
</tr>
<tr>
<td></td>
<td>(0.028)</td>
<td>(0.029)</td>
<td>(0.028)</td>
<td>(0.028)</td>
</tr>
<tr>
<td>Father's education a</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Junior high school</td>
<td>0.115**</td>
<td>0.026**</td>
<td>0.106**</td>
<td>-0.002</td>
</tr>
<tr>
<td></td>
<td>(0.016)</td>
<td>(0.010)</td>
<td>(0.017)</td>
<td>(0.016)</td>
</tr>
<tr>
<td>Senior high school</td>
<td>0.175**</td>
<td>0.052**</td>
<td>0.191**</td>
<td>-0.001</td>
</tr>
<tr>
<td></td>
<td>(0.020)</td>
<td>(0.020)</td>
<td>(0.032)</td>
<td>(0.027)</td>
</tr>
<tr>
<td>College</td>
<td>0.350**</td>
<td>0.033</td>
<td>0.354**</td>
<td>-0.062</td>
</tr>
<tr>
<td></td>
<td>(0.031)</td>
<td>(0.025)</td>
<td>(0.033)</td>
<td>(0.037)</td>
</tr>
<tr>
<td>Father's EGP class b</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Profs., managers (I,II)</td>
<td>0.070**</td>
<td>0.049*</td>
<td>0.080*</td>
<td>0.031</td>
</tr>
<tr>
<td></td>
<td>(0.025)</td>
<td>(0.019)</td>
<td>(0.035)</td>
<td>(0.030)</td>
</tr>
<tr>
<td>Routine nonmanal (III)</td>
<td>0.047</td>
<td>0.036</td>
<td>0.085</td>
<td>0.008</td>
</tr>
<tr>
<td></td>
<td>(0.025)</td>
<td>(0.025)</td>
<td>(0.045)</td>
<td>(0.031)</td>
</tr>
<tr>
<td>Small owners (IVa,IVb)</td>
<td>-0.057</td>
<td>0.028</td>
<td>0.063</td>
<td>-0.016</td>
</tr>
<tr>
<td></td>
<td>(0.035)</td>
<td>(0.033)</td>
<td>(0.037)</td>
<td>(0.027)</td>
</tr>
<tr>
<td>Foremen, skilled (V,VI)</td>
<td>-0.019</td>
<td>0.027</td>
<td>0.059</td>
<td>0.019</td>
</tr>
<tr>
<td></td>
<td>(0.031)</td>
<td>(0.018)</td>
<td>(0.031)</td>
<td>(0.029)</td>
</tr>
<tr>
<td>Semi- &amp; unskilled (VIIa)</td>
<td>-0.029</td>
<td>0.043*</td>
<td>-0.001</td>
<td>0.048</td>
</tr>
<tr>
<td></td>
<td>(0.026)</td>
<td>(0.017)</td>
<td>(0.039)</td>
<td>(0.031)</td>
</tr>
<tr>
<td>Number of siblings</td>
<td>-0.020**</td>
<td>0.004</td>
<td>-0.032*</td>
<td>0.013</td>
</tr>
<tr>
<td></td>
<td>(0.006)</td>
<td>(0.002)</td>
<td>(0.013)</td>
<td>(0.007)</td>
</tr>
<tr>
<td>Sex (male=1)</td>
<td>0.042*</td>
<td>0.031</td>
<td>0.076*</td>
<td>-0.020</td>
</tr>
<tr>
<td></td>
<td>(0.020)</td>
<td>(0.017)</td>
<td>(0.037)</td>
<td>(0.026)</td>
</tr>
<tr>
<td>Sex*Hukou origin</td>
<td>0.003</td>
<td>-0.014</td>
<td>-0.081</td>
<td>-0.013</td>
</tr>
<tr>
<td></td>
<td>(0.022)</td>
<td>(0.019)</td>
<td>(0.047)</td>
<td>(0.038)</td>
</tr>
</tbody>
</table>

N = 6996, 3140, 2937
Log-likelihood = -6169.4, -2881.4, -2471.4
Pseudo R-square = 0.104, 0.128, 0.145

Note: Robust standard errors in parentheses. Data are weighted. Within-province clustering effects are controlled. * p<.05; ** p<.01 (two-tailed test)

a Refer to “primary school or below”. b Refer to “farmers (IVc, VIIb)”. c Reported are Z-statistics testing coefficient differences across models.
In the meantime, the effect of hukou origin on transition to the vocational track is significant in the first period, and the size of the effect is lower than that for the academic track (0.079 vs. 0.202). The effect is also significant in the second period (about the same size), but the effect difference between the first two periods is insignificant (z=-0.31). In the third period, hukou origin has become a nonsignificant predictor for vocational school attainment, but the change of the effect is significant (z=2.95). The effects of father’s education and occupation on vocational school attainment have also become statistically insignificant in the second period. In the third period, having a father with senior high school or college education even negatively predicts the probability of going to the vocational track. These results imply that the vocational track has become a less favorable option to advantageous groups, particularly in the third period when college education is largely available. To more intuitively illustrate the empirical patterns, Figure 4.1a and 4.1b depict the effects of hukou origin on senior high school and college transitions in each year of transition.

The yearly effects are estimated based on the models presented in Table 4.4, and the logit coefficients of the hukou origin variable in the models are plotted to reflect the urban-rural gaps in educational opportunities. To indicate the trends of the logit effects, a kernel-weighted local polynomial smoothing procedure is used to smooth the non-linear curves (Fan and Gijbels 1994). Consistent with the results in Table 4.4, the curve in Figure 4.1a shows that the effect of hukou origin on senior high school transition has significantly increased in the period. The curve in Figure 4.1b further shows that the urban-rural gap in college opportunity has become significantly larger in the third period, suggesting increased educational inequalities in college attainment after the 1999 expansion.
Figure 4.1a: Trend of the Effect of Hukou Origin on Senior High School Transition

Figure 4.1b: Trend of the Effect of Hukou Origin on College Transition

Note: Dots in Figure 4.1a and 4.1b refer to the yearly estimated logit coefficients of hukou origin on senior high school and college transitions based on models in Table 4.4, respectively. The non-linear curves are smoothed by the kernel-weighted local polynomial smoothing procedure.
Explaining Increasing Educational Inequalities

Based on the general theoretical framework in Chapter 2, I interpret the trend reflected in Figure 4.1a as the result of relatively contracted senior high school opportunity in the 1990s. I also interpret Figure 4.1b through a structural and an institutional explanation (Hypothesis 3a and 3b). To further test these hypotheses, I estimate a set of hierarchical random coefficient models and examine whether the social origin effects on college attainment are significantly affected by measures of the “bottleneck” structure and rising college cost.

The level-1 independent variables are basically the same as those in Table 4.5, but the categorical measures of father’s education and occupation are substituted by the continuous measures of father’s year of schooling and ISEI score. This is to avoid over-dispersion in model estimation, as the number of the level-2 unit (year) is limited given the lack of educational fiscal statistical data before 1995. The level-2 variables include only two measures, a measure of the “bottleneck” structure (i.e. the difference in the transition rates to college and to senior high school) and a measure of rising college cost (i.e. average college tuition and fees). The correlation coefficient between the two measures equals 0.73 (p<.01). College premium is another possible predictor for level-2 effects. However, as discussed earlier, the theoretical implication of college premium for college attainment is not clear. The empirical results also report mixed findings about its effect.44 Therefore, to avoid theoretical and empirical complexities, I do not include a measure of college premium in the level-2 model. Table 4.7 presents the estimated fixed and random effects for five hierarchical random coefficient logit models.

---

44 Results are available upon request. One possible reason for the mixed findings is that the correlation coefficients between college premium and the other two measures are high: 0.88 with the “bottleneck” structure measure, 0.87 with the college cost measure. Both are significant at p<.01 level.
Table 4.7: Restricted Maximum Likelihood Estimation for Hierarchical Logit Random Coefficient Models on College Attainment, 1981-2010

<table>
<thead>
<tr>
<th>Fixed Effects</th>
<th>1981-2010</th>
<th>1995-2010</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Model 1</td>
<td>Model 2</td>
</tr>
<tr>
<td><strong>Level-1</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Hukou origin (urban=1)</em></td>
<td>0.527**</td>
<td>0.478**</td>
</tr>
<tr>
<td></td>
<td>(0.160)</td>
<td>(0.160)</td>
</tr>
<tr>
<td>Father’s years of schooling</td>
<td>0.067**</td>
<td>0.068**</td>
</tr>
<tr>
<td></td>
<td>(0.009)</td>
<td>(0.009)</td>
</tr>
<tr>
<td>Father’s ISEI</td>
<td>0.009**</td>
<td>0.015**</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>Number of siblings</td>
<td>-0.074*</td>
<td>-0.073*</td>
</tr>
<tr>
<td></td>
<td>(0.029)</td>
<td>(0.028)</td>
</tr>
<tr>
<td>Sex (male=1)</td>
<td>0.417**</td>
<td>0.417**</td>
</tr>
<tr>
<td></td>
<td>(0.118)</td>
<td>(0.118)</td>
</tr>
<tr>
<td>Sex*Hukou origin</td>
<td>-0.392*</td>
<td>-0.396*</td>
</tr>
<tr>
<td></td>
<td>(0.015)</td>
<td>(0.015)</td>
</tr>
<tr>
<td><strong>Slopes as outcomes</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hukou origin* C/H Rate difference</td>
<td>0.215*</td>
<td>0.243*</td>
</tr>
<tr>
<td></td>
<td>(0.034)</td>
<td>(0.042)</td>
</tr>
<tr>
<td>Father’s ISEI* C/H Rate difference</td>
<td>0.004**</td>
<td>(0.000)</td>
</tr>
<tr>
<td>Hukou* College tuition</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.087)</td>
<td>(0.086)</td>
</tr>
<tr>
<td>Father’s ISEI* College tuition</td>
<td></td>
<td>-0.000</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Random Effects</th>
<th>Estimates of Standard Deviations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Within individual residuals, $\mu_{ij}$</td>
<td>.988</td>
</tr>
<tr>
<td>Year intercept, $\delta_{ij}$</td>
<td>1.013**</td>
</tr>
<tr>
<td>Slope random effects</td>
<td></td>
</tr>
<tr>
<td><em>Hukou origin slope</em> C/H Rate difference</td>
<td>.378**</td>
</tr>
<tr>
<td>Father’s ISEI slope* C/H Rate difference</td>
<td>.011**</td>
</tr>
<tr>
<td>*Hukou origin slope * College tuition</td>
<td></td>
</tr>
<tr>
<td>Father’s ISEI slope * College tuition</td>
<td></td>
</tr>
<tr>
<td>*Hukou origin slope *(C/H Rate difference + College Tuition)</td>
<td></td>
</tr>
</tbody>
</table>

| N                      | 5560  | 5560  | 2568  | 2568  | 2568  |
| Level-2 unit (Year)    | 26    | 26    | 12    | 12    | 12    |
| Log-likelihood         | -3373.6 | -3382.1 | -3642.0 | -3543.4 | -3644.0 |

Note: Reported are estimated coefficients, with robust standard errors in parentheses. The indicator “C/H Rate difference” refers to the difference in the transition rates to college and to senior high school. It is the raw score (scaled to be 10 times larger). The scale of college tuition is divided by 1000. Father’s ISEI and year of schooling are centered at the group means. Data are weighted.

* p<.05; ** p<.01 (two-tailed test)
In all the models, the level-2 intercepts are allowed to vary. According to the effect difference tests in Table 4.5, I only estimate the random effects of hukou origin and father’s ISEI, and set the effects of the other level-1 variables fixed. The first two models examine how the slopes of hukou origin and father’s ISEI vary by different levels of constraint in China’s educational opportunity structure. Model 1 specifies that the slopes of hukou origin are determined by the “bottleneck” structure only, and vary by year. Model 2 further allows the slopes of father’s ISEI to vary by the constricted level of the bottleneck structure and by year. Similarly, Model 3 and 4 test whether college tuition significantly affects the slopes of hukou origin and father’s ISEI. Model 5 includes both measures of the “bottleneck” structure and college tuition in the level-2 random effect estimation, so that the effect of hukou origin on college attainment is now determined both by the educational opportunity structure and college cost.

The results provide strong support for Hypothesis 3a and 3b. Model 1 shows that the effect of hukou origin significantly differs by the level of constraint in the “bottleneck” structure. Model 2 further shows that the effect of father’s occupation is also significantly affected by the educational opportunity structure. Others being equal, a 10 percent increase in the difference of the college vs. senior high school transition rates would lead to a 0.24 unit increase in the coefficient of hukou origin and a 0.004 percent increase in the slope of father’s ISEI. The time period in Model 3 to 5 changes to 1995-2010 because of the inclusion of college tuition data. Model 3 indicates that college tuition is also a significant predictor of the slope of hukou origin. Net of other factors, every 1000 yuan increase in college tuition predicts a 0.348 unit increase in the coefficient of hukou origin. When this effect is controlled, Model 4 shows that college tuition does not significantly affect the slope of father’s ISEI. Finally, Model 5 suggests that both the “bottleneck” structure and college cost are significant determinants of hukou origin’s effect on
Figure 4.2a: Scatterplot between the Bottleneck Structure and *Hukou* Origin Effect on College Attainment, Fitted Values, over Years of Transition

Note: Figures 2a and 2b are graphed based on Model 2, Table 4.7. The measure of bottleneck structure is the difference between the transition rates to college and to senior high school.

Figure 4.2b: Fitted Slopes of Father’s ISEI on College Attainment, by Level of Constraint in the Bottleneck Structure
college attainment. Therefore, both the structural story (the “bottleneck” educational opportunity structure) and the institutional story (rising college cost) are confirmed by the multilevel modeling results.

The patterns are intuitively illustrated in Figure 4.2a, Figure 4.2b and Figure 4.3. Figure 4.2a plots the estimated hukou origin effect over the level of constraint in the “bottleneck” structure. The tendency is that higher hukou origin effect corresponds with higher levels of bottleneck constraint.

Figure 4.2b further shows that the fitted slope of father’s ISEI on college attainment is larger when the bottleneck constraint is higher. Figure 4.3 graphs the relationship between college tuition and hukou origin effect on college attainment, with a 95% confidence interval for the estimated effects. Similar to that in Figure 4.2a, the tendency is that higher hukou origin effect matches with higher college tuition.
CONCLUSION

In this chapter, I have examined the changing patterns of educational stratification during China’s economic reform-era from 1981 to 2010. Using the pooled data from five highly comparable national representative samples, I investigate changes in the social origin effects on educational transitions. I find that urban-rural differentials and social origin differentials in access to vocational school both decreased in the 1990s, but differentials in the transition to senior high school increased because of relatively contracted high school opportunity in this period. I also find that after the 1999 expansion, educational inequality in going to senior high school was generally declining (with a caveat that the effect differences are statistically insignificant), but the allocation of college opportunity became less equal. Urban-rural inequality in the transition to college increased, and father’s class status began to play a more important role in determining one’s likelihood of entering college. As a result, the rapid expansion of higher education in China since 1999 mainly benefited urban children and children from better-off families (e.g., managers/professionals).

In this context, unlike the “persistent inequality” pattern generally observed in industrial societies, reforming China underwent combined patterns of increasing inequality both in access to senior high school in the 1990s and to college after 1999, as well as decreasing inequality in access to vocational school in the 1980s and to senior high school since the late 1990s. In my assessment, the results signal that the advantaged social groups have shifted their previous advantages in the vocational track to the upper secondary academic track (senior high school) in the 1990s, and switched their advantages in the academic track to college since 1999. In this regard, educational inequality in reform-era China was maintained not only in quantity, but also in quality (Lucas 2001).
The combined patterns of educational inequalities can hardly be anticipated by the existing theoretical explanations, such as the rational choice theory and the “maximally maintained inequality” (MMI) hypothesis (Raftery and Hout 1993; Hout 2006). Based on the general theory proposed in Chapter 2, I argue that these patterns can be understood through the structural and institutional changes in China’s educational system. Particularly, China’s increasingly constricted “bottleneck” structure has significantly blocked children from disadvantaged background to continue college education. The rising educational cost has also posted new constraints on the disadvantaged groups for affording college. Results based on hierarchical linear modeling have provided strong support for both the structural and institutional perspectives. They show that both the level of constraint in the educational opportunity structure and the level of college cost are significantly positive determinants of the social origin effects on college attainment.

Therefore, educational expansion in China calls for attention to the relevance of educational opportunity structure in understanding schooling inequalities. Although China’s bottleneck educational opportunity structure is unique, it does provide general implications for understanding the process of educational stratification. Specifically, when there is a selection barrier in the educational system that systematically blocks educational opportunities for certain social groups, educational inequality at the next higher level would be likely to increase because of the uneven social selection in the earlier process. My formal modeling in Chapter 2 has illustrated the scope conditions for this to happen. In this regard, I argue that educational opportunity structure in a society is one important factor to explain the various empirical patterns of educational inequalities (see also Gerber 2000; Gerber and Hout 1995 for the Russian case).

My study also suggests that institutional change in the educational system is another
important dimension to understand the shifting patterns of changing educational inequalities. In the Chinese case, my analyses show that rising educational cost has significantly reduced the chances of the disadvantaged social groups to get ahead in educational attainment. If in another place we also observe enlarged educational cost, which implies larger resource differentials among social groups, then educational inequality is expected to increase. In Chapter 2, an illustrative simulation on how changes in the resource requirement of schools affect the allocation of educational opportunities is provided (Table 2.2, Chapter 2).

Moreover, as both the structural and institutional changes in the educational system are intentionally driven by the state, the Chinese case also reminds us that state dynamics can systematically change the mechanisms of allocating educational opportunities (Deng and Treiman 1997; Whyte and Parish 1984; Zhou and Hou 1999; Zhou et al 1996, 1997). In China, the state controls school enrollment though the national enrollment planning system, which decides not only the level of competition among students, but also the availability of schooling opportunities. Therefore, the state has played a significant role in the allocation process of educational opportunities in China. Comparing China with other societies in terms of educational planning and allocating processes will be a future step of this research.

Finally, my study suggests that certain social groups tend to secure educational advantages in steps. When higher education was only modestly available in the 1990s, their main advantage was on senior high school education; when college education was commonly available since 1999, they then shifted their advantage from senior high school to college. In this light, whereas China has a long tradition of merit-based educational system, China’s recent educational reforms have ironically created structural and institutional constraints that limited the upward mobility channel for disadvantaged groups through college education. Therefore, in the short run,
increased educational inequality can be transmitted into the labor market (see Chapter 2 for a formal illustration of this point). In the long-run, enlarged educational inequality can enhance intergenerational transmission of socioeconomic status, and the role of education as an important channel for socioeconomic upward mobility will be weakened (as observed in Ireland and Britain by Bukodi et al 2015; Breen and Goldthorpe 2001; Breen and Whelan 1993; Erikson and Goldthorpe 2010). In the next two empirical chapters, I will demonstrate the far-reaching consequences of the increasing educational inequality in China on the labor market returns to education.
Chapter 5


INTRODUCTION

The relationship between education and first occupation is an important issue in sociology. In the past years scholars have paid increasing attention to the cross-national differences in the association between education and first work (e.g. Allmendinger 1989; Andersen and van de Werfhorst 2010; Ishida et al 1995; Kerckhoff, 1995a, 2000, 2001; Maurice et al 1986; Mayer and Carroll 1987; Shavit and Muller 1998). Comparative studies on the school-to-work transition have shown that there are substantial cross-national variations in term of the impact of education on labor market entry outcomes. It is also widely accepted that this variation largely reflects differences in educational systems, especially with respect to the level of institutionalized vocational specificity (e.g., Arum and Shavit 1995; Breen 2005; Korpi et al 2003; Shavit and Muller 1998). For instance, stronger education effects on work outcomes, as measured by occupational status, occupational prestige, or social class, are consistently found more in countries with substantial vocational components to their education systems (e.g., Germany, the Netherlands and Switzerland) than in countries with limited vocational components (e.g., the USA and UK).

While previous studies have highlighted how nation-specific institutional features of educational system shape the association between education and first occupation (Gerber 2004), the fact that such association could be moderated by the expansion in the educational system has largely been overlooked. Particularly, although educational expansion has taken place in many societies, few studies have demonstrated how educational expansion affects the school-to-work
transition process (see the discussion of Arum and Hout 1998). The impact of educational expansion on the education-first occupation linkage, however, is by no means trivial.

First, educational expansion could modify the institutional contexts of the educational system itself. In many cases, nation-specific institutional features of its educational system have changed after the expansion (Soskice 1994; Walker and Zhu 2003, 2008). Second, as more and more educated workers entering labor market during the process of educational expansion, the matching process of education and occupation can be altered consequentially (Erikson and Jonsson 1998). Specifically, the composition of educational groups changes due to the influx of more college-educated workers into the labor market. In other words, the individuals that search for first jobs before and after the expansion tend to differ in significant ways. Ignoring the changing composition of educational groups will bias our estimates of the effect of education on first work.

In this chapter, I study the dynamics through which educational expansion changes the association between education and first occupation, paying special attention to the changing composition of skilled labor in the labor market. Using recent nationally representative survey data from reforming China, I examine the association between educational attainment and first occupation for individuals who had completed education and entered the labor market from 1981 onward to 2008.

Here, the case of reforming China proves to be an especially suitable case. In the last three decades, China’s educational system has experienced significant expansion at all levels. One sudden reform took place in the late 1990s, creating an increasingly larger body of college students since 1999. Therefore, I ask: within the Chinese context of educational expansion, is the association between educational attainment and first occupations strengthened or weakened?
How does the process of obtaining the first occupation change throughout the process of educational expansion? Specifically, do the effects of social origin factors on the education-first occupation association change in different reform periods? In this chapter, I provide empirical investigations of these questions by controlling for the changes introduced to the labor market labor composition.

In the following sections, I first briefly review the relevant literature. I then propose a series of hypotheses based on the theoretical model proposed in Chapter 2. I move on to show that there exist significant differences among different labor market entry groups before and after the expansion. Particularly, I focus on the directions of the change in the education-first occupation association across these groups. Furthermore, I provide an in-depth examination of the variation by sex and hukou origin in the association.

LITERATURE REVIEW

The Institutional Perspective on School-to-work Transition

In most societies, the educational system is one key mechanism for channeling young people to occupational positions in the process of social stratification (Allmendinger 1989). Education systems, nevertheless, differ in various societies in many aspects, e.g. levels of standardization, stratification and vocational specificity (Shavit and Muller 1998; Kerckhoff 2001). Standardization refers to the extent to which schools meet common national standards for budgets, examinations, and/or teacher training. Stratification refers to the level of tracking in secondary education. A secondary-level tracking system aims to ensure that individuals who are qualified for tertiary education and vocational training respectively are both able to obtain desirable positions in the labor market (Ainsworth and Roscigno 2005). Vocational specificity
indicates the extent to which particular educational tracks prepare students for occupational specific skills (Kerckhoff 2001; Muller et al 1998). For example, in some societies such as Germany, education is closely related to the workplace and vocational training prepares the students with specific skills. In France, in contrast, the educational system provides more general training and vocational skills are mainly obtained on the job. Education in French case is weakly linked to the labor market.

Because of the different structures of educational institutions across various countries, the effect of education on first occupation also differs across institutional settings (Kariya 1989, 1998). The neo-institutionalist and modernization theory generally predict a convergence of education’s effects on occupational attainment (Treiman 1970; Treiman and Yip 1989). Empirical results, however, suggest that the linkage between education and occupational outcomes is largely dependent on the specific institutional features of the educational system in a given society (Kerckhoff 1995a, 2001). Both the size and the shape of educational effects in the school-to-work transition process are different across countries. Such variations are mainly attributed to the differences in the characteristics of educational systems (Shavit and Muller 1998). Specifically, countries characterized by highly stratified education systems with strong vocational components, extensive tracking, and early selection into tracks tend to also exhibit a stronger relationship between education and occupation, when compared to countries with less vocational training and stratification (Saar et al 2008; Scherer 2005). In other words, the more stratified and vocationally specific the educational system is, the more transparent potential employees’ qualifications are to employers, which then leads to stronger matching between education and occupation.
A handful of studies have systematically explored the relationship between the education system and occupational outcomes with multi-country data. For example, Shavit and Muller (1998: Introduction) have examined 12 countries, predicting the effect of education in an individual country by a set of macro-level characteristics of its education system (e.g., level of standardization, stratification, and vocational specificity). Their results indicate that stratification and vocational specificity have significantly enhanced the relationship between education and occupation. Using a similar framework, Breen (2005) argues that strong educational signaling was most evident in countries characterized by extensive dual tracking systems, i.e. those with a high level of work-school based training apprenticeships. Scherer (2005) has further found stronger effects of tertiary and upper secondary vocational education on the school-to-work transition in Germany as compared to Italy and Britain. In all cases, the emphasis is placed on the institutional arrangement of an existing educational system.

While the institutional perspective is helpful for international comparisons, it has several limitations. First, it is difficult to rule out alternative explanations, as the institutional effect may be confounded by “longstanding traditions, cultural patterns, or structural economic conditions” (Gerber 2004: 244). By conducting cross-time comparison within a single country (i.e. Russia), Gerber’s (2004) study shows that institutional changes in Russia during 1970 to 2000 did not significantly impact the association between education and first occupation.

Second, the institutional perspective ignores the fact that macro-level sociopolitical institutions may be more influential for the educational system in some societies than in others. In China, for example, strong political interventions in the educational expansion process have created new dynamics in the allocations of resources and in the distribution of opportunities (Wu and Zhang 2010; Zhou and Hou 1999). Furthermore, such impacts generated by the state
interventions can be further transmitted into the origin-based inequality patterns in the labor market.

Third, the matching between education and destination (in term of occupational attainment) is not an isolated process. Previous patterns of educational attainment affect the subsequent composition of skilled labor in the labor market, and thus further impacts the labor market sorting processes (Granovetter 1988; Logan 1996). Therefore, without considering the interaction between macro-based institutions and micro-based labor market screening and matching processes, we cannot fully understand the linkage between school and work.

Finally, the institutional perspective does not consider the possible endogeneity within educational institutions per se. Particularly, institutional rules of the educational system can change along with the educational expansion processes, thus creating further different institutional contexts that subsequently allocate individuals to school and then to work. Moreover, as the labor market composition of skilled workers may shift after the expansion (e.g., Juhn et al 2005), workers may also face different labor market opportunity structure in the context of educational expansion.

Based on these considerations, in this chapter I study how educational expansion shapes the trends and patterns of school-to-work transition during China’s reform period. I will demonstrate that both the structural and institutional changes in the educational system have significantly impacted the linkage between education and first occupation under China’s educational expansion.
Finding a Job in China

Current existing studies on job attainment in the Chinese context can largely be separated into four camps. The first camp emphasizes the political screening process in the labor market. Here, results have shown that occupational attainment in China is characterized by a dual path model: both education and party membership serve as selection criteria when individuals try to obtain elite positions (Walder 1995, 2003; Walder et al 2000a, 2000b). To put it more concretely, although educational credentials are required for both administrative and professional elite positions, educational credentials alone are not enough. Party membership is also required for administrative and managerial positions of high prestige, considerable authority and material benefits (see also Bian et al 2001; Zhou 2001).

The second camp of studies adopts a social network perspective and stresses how social networks (guanxi) facilitate job search (Bian 1997, 2002; Bian et al 1998). In this conceptualization, it is social capital (or social resources) embedded in individuals’ social networks that helps occupational attainment (Bian et al 2005; Lin 1990). Guanxi network, therefore, becomes an important mechanism in understanding the transition from school to work.

The third camp shifts the attention from occupational attainment per se to work unit placement (Bian 1994; Walder 1986, 1992; Wu 2002; Zhou et al 1997). The idea behind this camp is that in China, individuals’ work unit status can be as important as, if not more important than, their occupations (Xie and Wu 2005). Furthermore, a large body of literature exists also on the job-change patterns between work units or between sectors (e.g., Zhou et al 1997).

Last but not the least, the fourth camp of research emphasizes the labor market sorting and matching process (Wu and Xie 2003; Xie and Wu 2005). As Wu and Xie (2003) have demonstrated, the sorting process of heterogeneous groups of workers into the labor market
explains the higher returns to education in the market sector. Recent studies also highlighted that workers with higher education tend to be matched with better work units and occupations (Guo and Wu 2009).

Although these studies from each of the four camps outlined have characterized the particular institutional features of the Chinese society, e.g. the Communist party system, the guanxi networks, and the work unit system, etc., they have yet to pay enough attention to the enduring effects of educational expansion on the education-occupation relationship. Furthermore, these studies do not fully capture the fact that status attainment in China is not solely a function of the macro-level sociopolitical institutional features, but is also shaped by the micro-level rational choice of individuals and families.

EDUATIONAL EXPANSION AND SCHOOL-TO-WORK TRANSITION IN REFORMING CHINA: HYPOTHESES

The institutional linkage between educational and occupational attainment was once close in China. From 1978 to 1987, the state took the full charge of college graduates’ employment (baofenpei), assigning them to a work unit based on its national economic planning each year. While the graduates need not worry about their post-graduation employment, they had no autonomy in deciding the job placement. Starting from 1988, a “two-way selection” system (shuangxiang xuanze) was introduced so that both the graduates and the employers had a certain level of autonomy in the matching process. The college recommended willing students to employers, and employers finalized hiring based on the quality of the recommended students. This “two-way selection” system lasted for the next ten years. In 1998, along with the college expansion started in the late 1990s, the employment of college graduates became completely
marketized. Since then, together with the 1999 massive college expansion, students now have full autonomy in finding and obtaining jobs. In other words, these new cohorts of students are no longer assigned to jobs by the state or recommended to employers by the schools. Along with the sharp increase in the supply of college graduates, the unemployment rate for college graduates reached a high level. (Cai and Wang 2010). The average registered college employment rate was only about 70 percent in most of the years in the last decade (Feng et al 2009), leaving a huge population of unemployed college graduates. In 2009, it was estimated that over 1.9 million college students who graduated in that year would not be able to find a job.

In this sense, educational expansion has significantly changed the labor market opportunity structure for the post-expansion cohorts. First, the massive supply of college graduates after the expansion has tightened the labor market in which the supply of skilled labor is largely greater than the demand. Second, the declining cohort quality and the declining quality of colleges (see Figure 6.2, Chapter 6) have reduced the signaling value of college credentials in the labor market. Third, because of China’s “bottleneck” structure, the process of educational attainment has pushed a relatively larger proportion of college graduates that are also from better-off families into labor market. To reiterate, the labor market entry cohorts after the 1999 expansion are characterized by three distinctive characteristics, i.e. larger quantity, lower quality, and higher origin-education association than before.

The first two characteristics of the after-expansion cohorts would predict that the education-first occupation association may have become lower in the last decade. Given that the relative demand does not shift much in the short time period studied, an increasing supply of college graduates suggests that the competition for better jobs would be higher than before. Thus, only a relatively smaller proportion of college graduates are now able to seize the “good” job
opportunity, which is more sorting by quality structure of college. A lower cohort quality further implies that the value signaled by the college credential is declining. To put it differently, a certain proportion of the post-expansion college graduates are only comparable to the pre-expansion senior high school graduates in term of cohort quality. Therefore, the jobs that they are able to obtain may be less prestigious, which further exerts downward pressure on the income return to college. Taken together, I hypothesize that:

Hypothesis 1: The association parameter between college and first occupation will be lower after the 1999 expansion.

The second and third characteristics of the after-expansion cohorts further suggest that the college-first occupation association may vary across different social origin groups. The social origin composition in educational attainment may suggest that the signals of college credential and non-college credential could work differently for the advantaged vs. disadvantaged social groups. Specifically, given that individuals from better-off families have largely benefitted from college expansion, in this way, not having a college credential would signal low ability for children from advantageous social origins. However, such signaling mechanism may not hold true for children from worse-off families. Instead, having a college credential may even signal high ability for children from disadvantageous social groups. Here, if this story of differential signaling holds, we would then observe a larger college vs. senior high school opportunity differentials for children of advantaged background.

In addition, the increasing origin-education association suggests that if the structural change in the composition of college-educated labor favors the advantaged group, quantitatively speaking, the linkage between college education and job positions also becomes larger for the
advantageous group than for the disadvantaged group if the market is competitive and provides sufficient employment opportunity.\textsuperscript{45}

Compared to the disadvantaged group, the advantaged group now enjoys even larger advantages in using education for social mobility. If the effect of ascribed factors on occupational attainment does not shrink over time and post-expansion, we should then expect that across time, the advantaged group will be even better-positioned to obtain more desirable position using educational credentials. The implication behind is that educational expansion reduces the incentives for the disadvantaged group to maximize educational attainment for social mobility, but strengthens the incentives for the advantaged group. The mobility strategies of the advantaged group and the disadvantaged group thus will continue to diverge along the process of education-based social stratification.

In the meanwhile, the institutional linkage of the educational system to the labor market is generally declining in China. As mentioned above, previously colleges and universities could directly recommend, if not send, students to various firms and governmental institutions; now students need to instead find jobs by their own. The state also changed its role from “taking charges of students’ jobs” (\textit{baofenpei}) to “none of my business”, because it wants the labor market to be more efficient and competitive.

Combined with the fact that there is a tightening labor market in recent years, the importance of ascribed factors (such as parents’ social networks) in the job-searching process may grow. When finding jobs, children from the advantaged group can still depend on ascribed factors when their credential does not work, but the disadvantaged group is again in a more

\textsuperscript{45} Suppose the labor market is competitive and provides sufficient employment opportunity. If there are larger proportions of college graduates both from the higher class (say, from $p_{h1}$ to $p_{h2}$) and the lower class (say, from $p_{l1}$ to $p_{l2}$) than before, but the ratio of $p_{h1}$ over $p_{l1}$ is lower than $p_{h2}$ over $p_{l2}$, then when all the college graduates find jobs in the labor market, class-based proportions do not change in the job positions and the later ratio is still larger than the earlier one.
disadvantageous position in this aspect because they have fewer social resources (Bian 1994; Lin and Bian 1991). Therefore, a tightening labor market can accelerate class differentials in the school-to-work transition.

Hypothesis 2: On average, social origin differentials on the college-first occupation association will be larger after the 1999 expansion.

Together with the increasing importance of education as a means of resource allocation under market transition (Nee 1996), it is likely that children from the advantaged group have relatively more ways to access a better job across time because they have more advantages both in achieved and ascribed factors in the labor market. Class differentials in the education-destination association therefore tend to be increasing.

DATA, VARIABLES AND METHODS

The data used for this chapter were draw from the 1996 “Life History and Social Change” Survey and the 2005, 2006 and 2008 waves of China General Social Survey (CGSS). The currently available CGSS 2010 Survey does not have respondents’ first job information and is therefore omitted in the analyses for this chapter. These datasets provide national representative samples for both urban and rural residents. Data were pooled to obtain a larger sample size for adequate analytical leverage, especially for the most recent birth cohorts that were affected directly by educational expansion. Statistical tests using log-linear modeling have shown that there is no significant difference in the results across the sources of data (results available upon request).

Following the analytical strategies in the existing literature in the field, I have limited my analyses to people aged 14-35 at the time of first job attainment. I have distinguished three labor
cohorts based on the entry year of first occupation: 1981-1992, 1993-2002, and 2003-2008. The year 1981 is chosen to be the starting point so as to include the cohorts that entered college in the late 1970s. The year of 1992 witnessed Deng Xiaoping’s political tour to southern China, after which China deepened its economic reforms and marketization. The year 2003 marked the first year of labor market entry for those who went to (four-year) university in 1999, the year of the start of educational expansion. As suggested in Chapter 3, the number of newly graduated college students has increased yearly ever since 2003. To demonstrate the effect of the massive educational expansion starting in 1999, I have also constructed a variable indicating college entry year. Particularly, I have distinguished those who enter college before from those who entered after educational expansion (i.e. college entry year on/before and after 1998).

The dependent variables include respondent’s first occupation. First occupation is measured by a continuous ISEI score (Ganzeboom et al 1992) as well as a six-category EGP class schema (Wu and Treiman 2007), both commonly used in comparative stratification research. Specifically, following Wu and Treiman (2007), I have coded respondents’ first occupation into a 6-category version of the EGP scheme (Erikson, Goldthorpe, and Potocarero 1979; Ganzeboom, Liujkx, and Treiman 1989) for the Chinese case. The relationship between the 10-category version proposed by Erikson et al (1979) and the 6-category version used here is as shown in the following table (reprinted here from Table 4.1, Chapter 4).

Independent variables include education and party membership at the time of first job, *hukou* origin when aged 14 or 18 (depending on data availability; party membership using the party entry year information), and father’s education and occupation when aged 14 or 18. Following the standard procedure in the literature, I have used the CASMIN educational schema to operationalize educational attainment. The result is a five-level categorical variable: junior high
<table>
<thead>
<tr>
<th>Original classification</th>
<th>New classification</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. Large proprietors, higher professionals and managers</td>
<td>6</td>
</tr>
<tr>
<td>II. Lower professionals and managers</td>
<td>6</td>
</tr>
<tr>
<td>III. Routine nonmanual workers</td>
<td>5</td>
</tr>
<tr>
<td>IVa. Small proprietors with employees</td>
<td>4</td>
</tr>
<tr>
<td>IVb. Small proprietors without employees</td>
<td>4</td>
</tr>
<tr>
<td>V. Lower grade technicians and manual supervisors</td>
<td>3</td>
</tr>
<tr>
<td>VI. Skilled manual workers</td>
<td>3</td>
</tr>
<tr>
<td>VIIa. Unskilled and semiskilled manual workers</td>
<td>2</td>
</tr>
<tr>
<td>IVc: Self-employed farmers</td>
<td>1</td>
</tr>
<tr>
<td>VIIb. (Unskilled) agricultural workers</td>
<td>1</td>
</tr>
</tbody>
</table>


school or below (1a and 1b), senior high school (2b), vocational school (2a), three-year junior college (dazhuan) and four-year university (3a and 3b). Party membership is a dummy variable (yes=1). Hukou origin refers to respondents’ hukou status when aged 14 or 18 depending on the availability of this information in the CGSS data and is coded as a dummy variable (urban=1). Father’s occupational status when the respondent aged 14 or 18 is also measured by a continuous ISEI score and a categorical EGP class schema. The same procedure was carried out for respondents’ first job status. Additional control variables include sex, age at the time of first job, and a set of provincial dummies. To empirically measure the association between educational attainment and first job attainment, as well as any possible changes in the magnitude and/or pattern of said association by period, following conventional research on mobility tables, I rely on multinomial logistic modeling to investigate the relationship education and first occupation attainment under educational expansion.

Table 5.1 presents the summary statistics of the dependent and independent variables for the three labor market entry cohorts. It is evident that after educational expansion, the proportion of workers with college education has largely increased: for the 1993-2002 labor entry cohorts, only 6.78 percent of the total workers had graduated from four-year colleges. The figure jumped to
Table 5.1: Summary Statistics for Dependent and Independent Variables

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>First job ISEI</td>
<td>37.1</td>
<td>34.4</td>
<td>39.8</td>
<td>44.0</td>
</tr>
<tr>
<td></td>
<td>(14.95)</td>
<td>(14.07)</td>
<td>(15.30)</td>
<td>(14.99)</td>
</tr>
<tr>
<td>Year of education</td>
<td>9.9</td>
<td>9.1</td>
<td>10.5</td>
<td>11.9</td>
</tr>
<tr>
<td>Educational levels %</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Junior high or below</td>
<td>56.45</td>
<td>65.40</td>
<td>47.41</td>
<td>34.04</td>
</tr>
<tr>
<td>Senior high</td>
<td>17.43</td>
<td>19.02</td>
<td>15.73</td>
<td>13.71</td>
</tr>
<tr>
<td>Vocational</td>
<td>11.93</td>
<td>8.19</td>
<td>17.96</td>
<td>14.53</td>
</tr>
<tr>
<td>Junior college</td>
<td>8.51</td>
<td>4.75</td>
<td>12.12</td>
<td>18.49</td>
</tr>
<tr>
<td>University</td>
<td>5.68</td>
<td>2.64</td>
<td>6.78</td>
<td>19.24</td>
</tr>
<tr>
<td>Party membership</td>
<td>.025</td>
<td>.018</td>
<td>.030</td>
<td>.053</td>
</tr>
<tr>
<td>Hukou origin</td>
<td>.376</td>
<td>.353</td>
<td>.390</td>
<td>.462</td>
</tr>
<tr>
<td>Father’s ISEI</td>
<td>32.58</td>
<td>32.17</td>
<td>32.65</td>
<td>34.71</td>
</tr>
<tr>
<td></td>
<td>(15.30)</td>
<td>(15.18)</td>
<td>(15.21)</td>
<td>(16.06)</td>
</tr>
<tr>
<td>Age</td>
<td>19.6</td>
<td>18.3</td>
<td>20.8</td>
<td>22.9</td>
</tr>
<tr>
<td>Sex (%)</td>
<td>48.2</td>
<td>48.1</td>
<td>48.3</td>
<td>48.2</td>
</tr>
<tr>
<td>N</td>
<td>13957</td>
<td>8107</td>
<td>4383</td>
<td>1467</td>
</tr>
</tbody>
</table>

Note: Data are weighted. Standard deviations in parentheses.

19.24 percent for the 2003-2008 labor market entry cohorts. Moreover, in 2003-2008, there is also a larger percentage of people with urban hukou origin and party affiliation. Finally, father’s average ISEI is also slightly higher for the 2003-2008 labor entry cohorts than for the earlier cohorts. These figures further demonstrated that the labor entry cohorts are different before and post educational expansion along a variety of dimensions.

RESULTS

The major analytical objective is to assess the extent of association between educational attainment and labor-market positions at entry. To describe the distribution of first occupation by education, Table 5.2 presents the outflow mobility table by education at the time of first job for all the labor market entry cohorts. According to Table 5.2, college educated workers now had a lower percentage of ending up with a professional or managerial occupation as the first job. In
1981-1992, over half of the college graduates obtained a professional or managerial first job, and the figure went down to 45.44 percent in 1993-2002 and further to 34.42 percent in 2003-2008.

Table 5.3 further presents the OLS estimates. The first three columns use father’s ISEI when aged 14 or 18 as an indicator of social origin status, whereas the last three columns use father’s EGP class. Comparing the models, we find that the linkage between high education and a good first job is declining across cohorts. In 1981-1992, net of other factors, college-educated workers’ first job ISEI on average was 26.6 points higher than those with compulsory education or below, but the advantage declined to 24.5 point in the second period and to 19.8 point in 2003-2008. At the same time, the effect of communist party affiliation is declining, too. After 2002, first job ISEI differed only insignificantly between party members and non-party members.

In contrast, the effects of hukou origin and father’s origin status tend to be increased in the latest period. Controlling for others, workers with urban hukou origin enjoyed 2.5-point higher first job ISEI than those with rural hukou origin in 2003-2008, whereas in 1993-2002, the advantage was only 1.5 point higher. Similarly, net of others, one point increase in father’s ISEI
## Table 5.3: OLS Estimates of the ISEI score of First Job, 1981-2008

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>Model 1</th>
<th></th>
<th>Model 2</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex</td>
<td>0.198</td>
<td>-1.036</td>
<td>-0.597</td>
<td>0.126</td>
<td>-1.057</td>
</tr>
<tr>
<td></td>
<td>(0.317)</td>
<td>(0.552)</td>
<td>(0.695)</td>
<td>(0.310)</td>
<td>(0.569)</td>
</tr>
<tr>
<td>Age</td>
<td>0.120**</td>
<td>0.141**</td>
<td>-0.015</td>
<td>0.120**</td>
<td>0.151**</td>
</tr>
<tr>
<td></td>
<td>(0.025)</td>
<td>(0.042)</td>
<td>(0.091)</td>
<td>(0.027)</td>
<td>(0.043)</td>
</tr>
<tr>
<td>Education</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Senior high</td>
<td>7.429**</td>
<td>6.059**</td>
<td>6.121**</td>
<td>7.552**</td>
<td>6.127**</td>
</tr>
<tr>
<td></td>
<td>(0.554)</td>
<td>(0.724)</td>
<td>(1.313)</td>
<td>(0.561)</td>
<td>(0.691)</td>
</tr>
<tr>
<td>Vocational</td>
<td>13.352**</td>
<td>10.075**</td>
<td>4.850**</td>
<td>13.629**</td>
<td>10.288**</td>
</tr>
<tr>
<td></td>
<td>(1.160)</td>
<td>(0.777)</td>
<td>(1.174)</td>
<td>(1.152)</td>
<td>(0.791)</td>
</tr>
<tr>
<td>Junior college</td>
<td>20.831**</td>
<td>17.975**</td>
<td>13.599**</td>
<td>21.036**</td>
<td>18.084**</td>
</tr>
<tr>
<td></td>
<td>(1.265)</td>
<td>(0.994)</td>
<td>(1.177)</td>
<td>(1.220)</td>
<td>(1.002)</td>
</tr>
<tr>
<td></td>
<td>(1.640)</td>
<td>(1.345)</td>
<td>(2.520)</td>
<td>(1.607)</td>
<td>(1.365)</td>
</tr>
<tr>
<td>Party membership</td>
<td>8.008**</td>
<td>5.160**</td>
<td>1.791</td>
<td>8.053**</td>
<td>4.998*</td>
</tr>
<tr>
<td></td>
<td>(1.790)</td>
<td>(1.838)</td>
<td>(1.661)</td>
<td>(1.770)</td>
<td>(1.814)</td>
</tr>
<tr>
<td>Hukou origin</td>
<td>2.814**</td>
<td>1.606*</td>
<td>2.521**</td>
<td>3.205**</td>
<td>1.589*</td>
</tr>
<tr>
<td></td>
<td>(0.565)</td>
<td>(0.668)</td>
<td>(0.487)</td>
<td>(0.695)</td>
<td>(0.733)</td>
</tr>
<tr>
<td>Father’s ISEI</td>
<td>0.103**</td>
<td>0.085**</td>
<td>0.115**</td>
<td>0.104**</td>
<td>0.085**</td>
</tr>
<tr>
<td></td>
<td>(0.011)</td>
<td>(0.027)</td>
<td>(0.039)</td>
<td>(0.011)</td>
<td>(0.027)</td>
</tr>
</tbody>
</table>

Father’s class\(^b\)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Profs., managers (I,II)</td>
<td>3.125**</td>
<td>3.695**</td>
<td>5.747*</td>
</tr>
<tr>
<td></td>
<td>(0.708)</td>
<td>(1.317)</td>
<td>(2.231)</td>
</tr>
<tr>
<td>Routine nonmanual (III)</td>
<td>3.068**</td>
<td>1.992</td>
<td>2.390</td>
</tr>
<tr>
<td></td>
<td>(0.840)</td>
<td>(1.259)</td>
<td>(1.552)</td>
</tr>
<tr>
<td>Small owners (IVa,IVb)</td>
<td>3.208**</td>
<td>2.335*</td>
<td>1.709</td>
</tr>
<tr>
<td></td>
<td>(0.811)</td>
<td>(0.918)</td>
<td>(1.522)</td>
</tr>
<tr>
<td>Foremen, skilled (V,VI)</td>
<td>0.083</td>
<td>-0.168</td>
<td>0.575</td>
</tr>
<tr>
<td></td>
<td>(0.674)</td>
<td>(0.946)</td>
<td>(1.396)</td>
</tr>
<tr>
<td>Semi- &amp; unskilled (VIIa)</td>
<td>-0.108</td>
<td>0.449</td>
<td>3.917</td>
</tr>
<tr>
<td></td>
<td>(0.614)</td>
<td>(0.922)</td>
<td>(2.543)</td>
</tr>
</tbody>
</table>

Constant | 20.988** | 25.963** | 32.350** | 23.526** | 27.641** | 34.533** |
|          | (1.137) | (1.734) | (2.683) | (1.105) | (1.481) | (2.351) |

| Observations | 7,695 | 4,073 | 1,324 | 7,732 | 4,095 | 1,345 |
| R-squared    | 0.332 | 0.310 | 0.310 | 0.329 | 0.313 | 0.309 |

Note: Robust standard errors in parentheses. Data are weighted. Within-province clustering effects are controlled.

* p<.05; ** p<.01 (two-tailed test)
\(^a\) Refer to “Below senior high”.
\(^b\) Refer to “farmers (IVc, VIIb)”.

161
Table 5.4: Multinomial Logit Estimates of First Job Entry Class (reference=VIIb)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>I II</td>
<td>III</td>
<td>IV</td>
<td>V VI</td>
<td>VIIa</td>
<td>III</td>
<td>IV VI</td>
<td>VIIa</td>
</tr>
<tr>
<td>Sex</td>
<td>0.137</td>
<td>-0.421**</td>
<td>0.336**</td>
<td>0.382**</td>
<td>0.414**</td>
<td>-0.208</td>
<td>-0.659**</td>
<td>0.037</td>
</tr>
<tr>
<td></td>
<td>(0.104)</td>
<td>(0.142)</td>
<td>(0.116)</td>
<td>(0.100)</td>
<td>(0.114)</td>
<td>(0.224)</td>
<td>(0.210)</td>
<td>(0.140)</td>
</tr>
<tr>
<td>Age</td>
<td>0.067**</td>
<td>0.066**</td>
<td>0.027*</td>
<td>0.044**</td>
<td>0.056*</td>
<td>0.070**</td>
<td>0.067**</td>
<td>0.013</td>
</tr>
<tr>
<td></td>
<td>(0.013)</td>
<td>(0.013)</td>
<td>(0.011)</td>
<td>(0.010)</td>
<td>(0.011)</td>
<td>(0.012)</td>
<td>(0.012)</td>
<td>(0.012)</td>
</tr>
<tr>
<td>Education</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Senior high</td>
<td>2.310**</td>
<td>2.074**</td>
<td>0.690**</td>
<td>2.102**</td>
<td>1.020**</td>
<td>2.284**</td>
<td>2.631**</td>
<td>0.842**</td>
</tr>
<tr>
<td></td>
<td>(0.167)</td>
<td>(0.221)</td>
<td>(0.173)</td>
<td>(0.212)</td>
<td>(0.200)</td>
<td>(0.423)</td>
<td>(0.288)</td>
<td>(0.293)</td>
</tr>
<tr>
<td>Vocational</td>
<td>4.338**</td>
<td>3.388**</td>
<td>1.554**</td>
<td>2.198**</td>
<td>2.085**</td>
<td>3.630**</td>
<td>2.654**</td>
<td>1.183**</td>
</tr>
<tr>
<td></td>
<td>(0.366)</td>
<td>(0.283)</td>
<td>(0.287)</td>
<td>(0.279)</td>
<td>(0.299)</td>
<td>(0.284)</td>
<td>(0.303)</td>
<td>(0.211)</td>
</tr>
<tr>
<td>Junior college</td>
<td>5.385**</td>
<td>4.233**</td>
<td>1.855**</td>
<td>2.150**</td>
<td>3.526**</td>
<td>3.845**</td>
<td>1.932**</td>
<td>1.972**</td>
</tr>
<tr>
<td></td>
<td>(0.674)</td>
<td>(0.681)</td>
<td>(0.630)</td>
<td>(0.770)</td>
<td>(0.617)</td>
<td>(0.472)</td>
<td>(0.451)</td>
<td>(0.438)</td>
</tr>
<tr>
<td>College</td>
<td>7.419**</td>
<td>5.600**</td>
<td>2.641*</td>
<td>3.678**</td>
<td>2.561*</td>
<td>7.991**</td>
<td>6.160**</td>
<td>3.564**</td>
</tr>
<tr>
<td></td>
<td>(1.028)</td>
<td>(1.091)</td>
<td>(1.105)</td>
<td>(1.082)</td>
<td>(1.005)</td>
<td>(0.954)</td>
<td>(1.011)</td>
<td>(0.944)</td>
</tr>
<tr>
<td>Party membership</td>
<td>2.261**</td>
<td>2.043**</td>
<td>0.970</td>
<td>1.415*</td>
<td>1.102</td>
<td>1.014</td>
<td>1.439**</td>
<td>0.215</td>
</tr>
<tr>
<td></td>
<td>(0.647)</td>
<td>(0.711)</td>
<td>(0.614)</td>
<td>(0.598)</td>
<td>(0.586)</td>
<td>(0.550)</td>
<td>(0.453)</td>
<td>(0.499)</td>
</tr>
<tr>
<td>Hukou origin</td>
<td>2.825**</td>
<td>3.653**</td>
<td>2.948**</td>
<td>3.584**</td>
<td>3.978**</td>
<td>3.517**</td>
<td>3.842**</td>
<td>1.962**</td>
</tr>
<tr>
<td></td>
<td>(0.357)</td>
<td>(0.226)</td>
<td>(0.254)</td>
<td>(0.242)</td>
<td>(0.238)</td>
<td>(0.354)</td>
<td>(0.374)</td>
<td>(0.276)</td>
</tr>
<tr>
<td>Father’s IS/EI</td>
<td>0.040**</td>
<td>0.041**</td>
<td>0.019**</td>
<td>0.028**</td>
<td>0.018**</td>
<td>0.036**</td>
<td>0.038**</td>
<td>0.029**</td>
</tr>
<tr>
<td></td>
<td>(0.004)</td>
<td>(0.005)</td>
<td>(0.005)</td>
<td>(0.006)</td>
<td>(0.006)</td>
<td>(0.010)</td>
<td>(0.001)</td>
<td>(0.001)</td>
</tr>
<tr>
<td></td>
<td>(0.513)</td>
<td>(0.548)</td>
<td>(0.523)</td>
<td>(0.480)</td>
<td>(0.489)</td>
<td>(0.427)</td>
<td>(0.631)</td>
<td>(0.447)</td>
</tr>
<tr>
<td>Observations</td>
<td>7,808</td>
<td>7,808</td>
<td>7,808</td>
<td>7,808</td>
<td>7,808</td>
<td>4,186</td>
<td>4,186</td>
<td>4,186</td>
</tr>
</tbody>
</table>

Note: Robust standard errors in parentheses. Data are weighted. Within-province clustering effects are controlled. * p<.05; ** p<.01 (two-tailed test)
could lead to 0.085 point increase in first job ISEI in the second period and 0.115 point increase in the third. Moreover, the last three columns suggest that the advantage of father being a manager or professional on first job attainment tends to be the largest over in the studied period. These patterns are consistent if first job status is measured categorically. Table 5.4 presents the results of the determinants on the allocation of first job class position. Multinomial logit regressions have been used to estimate the effects of educational qualifications and social origin backgrounds on the likelihood of entry into a certain class. Class VIIb (agricultural workers) and being in the agricultural sector are used as the reference groups. Again, the results in Table 5.4 reveal that highly qualified people are advantaged in access to professional and managerial positions. The higher the level of educational qualification attained, the greater the chance of being recruited into the service class (I, II). However, the advantage of college educated workers seems to be declined since 2003. In addition, the effect of party membership disappeared in the latest period. Similar to the results in Table 5.3, the effects of hukou origin and father’s social origin status share a U-shape trend: the magnitude of the coefficients tend to be first declining and then increasing. Particularly, both Table 5.3 and 5.4 results indicate that after educational expansion, social origin factors have become more important than before.

CONCLUSION

Most recent comparative research on the school-to-work transition focuses on the roles of institutional arrangements in the stratification processes mainly in industrial societies (Kerckhoff 2001). Different from most cases in the literature, reforming China has been an industrializing country that has some distinctive characteristics in shaping the transition from school to work. One notable feature is that China has experienced significant educational expansion in the past
three decades. More importantly, starting from 1999, educational reforms in China have introduced both institutional and structural changes in the educational system that have subsequently affected the linkage between education and first work in the labor market.

In this chapter, I have concentrated on the experience of reforming China and investigated the association between education and first work under the context of educational expansion. By using multiple measurements of labor market entry positions, the empirical analyses suggest that China’s stratified educational system functions as a sorting and allocating mechanism. Particularly, higher educational credentials confer a higher return of occupational status. Moreover, secondary vocational education seems to give individuals a higher return of status in first jobs than senior high school education. In addition, compared to the 1980s, father’s occupation was a more important determinant for children’s occupational attainment in the 1990s, controlling for education.

The analyses also show that there are significant differences among those who entered the labor market before and after massive educational expansion in the late 1990s. After demonstrating the differences in cohort quality, the results show that there are still significant changes in the returns to first job occupational status for comparable cohorts. Without considering changes in the composition of skilled labor in the labor market introduced by educational expansion, our estimates of the effects of education on labor market entry positions may to be biased. A further step of the project is to identify the causal effects of education on first job socioeconomic status.
Chapter 6
Who Falls Behind in the College Premium? Heterogeneity, Selection and the Policy Effect of Educational Expansion on Earnings in Reforming China 1981-2010

INTRODUCTION

The linkage between college enrollment and changing returns to education has been well-documented in recent years (Card 2001; Carneiro and Lee 2007, 2011; Fortin 2006; Hendel et al 2005; Goldin and Katz 1999; Katz and Autor 1999). In the context of educational expansion, scholars find that the shifting supply of college graduates in the economy is intertwined with the trend of earnings inequality, particularly the trend of the college premium (Amin 2011; Devereux et al 2011; Gosling et al 2000; Katz and Murphy 1992; Walker and Zhu 2008). On one hand, the increasing supply of skilled workers in the labor market, together with the labor market demand, shifts the price of the college credential (Autor et al 1998, 2008; Murphy and Welch 1992, 1993); on the other hand, the increasing share of college graduates in the labor market cohorts enlarges the ability (and/or skills) heterogeneity among the college-educated, and thus changes the value associated with college credential (Card and Lemieux 2001; Hoxby and Terry 1999; Juhn et al 2005; Rosenbaum 2003; Taber 2001; Zheng 2009). Therefore, both the quantity and quality of college graduates have changed under educational expansion, which corresponds to a price effect and a composition effect, respectively (Carneiro and Lee 2007, 2011; Lemieux 2006).

Thus far, much research has focused on disentangling the two effects so as to obtain an accurate estimation of the quality-adjusted college premium (e.g., Carneiro and Lee 2007, 2009, 2011; Heckman and Vytlačil 1998; Juhn et al 2005; Kurtzon 2012; Moretti 2004; Taber 2001; Rosenbaum 2003). Much less attention has been placed on understanding the social processes...
through which educational expansion introduces these effects and how such effects further vary across different social groups (Brand and Xie 2010; Carneiro and Lee 2007; Deschenes 2007; Hout 2012; Strayer 2002; Xie et al 2012). Particularly, given that schools provide labor forces for the labor market, the quantity and composition of the college graduates supplied are determined by the earlier social selection processes of college attainment under educational expansion (Balart 2009; Carneiro and Lee 2007; Heckman and Vytlacil 2005; Zheng 2009). Since educational attainment is characterized by a sequence of social inclusion and exclusion of children with different social backgrounds (Ishida et al 1995; Kerckhoff 1995a, 2001; Muller and Karle 1993), increasing the supply of college graduates would increase not only the heterogeneity in ability (or aptitude), but also the heterogeneity in social backgrounds (see also Figure 2.1, Chapter 2).

Except for a few studies (e.g. Attewell and Lavin 2007; Brand 2010; Brand and Davis 2011; Brand and Xie 2010; Deschenes 2007; Tsai and Xie 2008, 2011), the increasing variation in social backgrounds for college graduates and its implications are not well surveyed in the current literature. One reason is that most of the studies on this topic are in the tradition of labor economics, which typically assumes that the return to schooling is “constant across the population or is a single random variable” (Deschenes 2007: 265). However, as Brand and Xie (2010) have pointed out, the treatment effect of college on earnings is essentially heterogeneous (see also Carneiro and Lee 2007; Carneiro et al 2010; Heckman and Vytlacil 2005; Heckman et al 2006). Ignoring such heterogeneity of the college premium among different social origin

46 See also Barrow and Rouse (2005) on the different returns to education by race and ethnicity, Card and Lemieux (1994, 2007) on gender and race, and Averett and Burton (1996) on gender.

47 In recent years, Heckman and his colleagues have published a series of papers focusing on the heterogeneity of treatment effects (e.g., Carneiro et al 2010, 2011; Heckman 2001, 2005; Heckman and Sedlacek 1985; Heckman and
groups, we miss the cumulative social process (DiPrete and Eirich 2006; Jencks and Tach 2006) through which educational expansion generates earnings inequality structure in the labor market.

In addition, the changing composition of social background of college graduates suggests that the social groupings of the college-educated are different over time. In sociology, there is a rich body of literature on how social capital (or social resources) helps people to be better awarded in the labor market (Bian 1997; Bian et al 2015; Coleman 1988; Granovetter 1985; Lin 1999). If a larger share of college students comes from advantageous social groups, more would be able to use their better-off social background to compensate for any devaluation of college credentials in the labor market. Therefore, not only are the demographic characteristics (beyond ability) of the college graduates shifting across cohorts, but the labor market sorting and matching processes also differ over time (Granovetter 1988). Without considering the important role of social background heterogeneity in the linkage between educational expansion and the college premium, we cannot fully understand the sources and mechanisms underlying the changing returns to education.  

In this chapter, I fill in the gap through examining the social origin heterogeneity in the college premium in the context of educational expansion. Using the case of reforming China, I ask the following two questions: first, how do the patterns of earnings inequality change under educational expansion? Second, which group has benefited more or suffered less in terms of the college premium under educational expansion? By answering these questions, I aim to provide not only a solid depiction of China’s labor market inequality structure in the reform period, but

---

48 It is possible that ability is positively correlated with social background if we view ability not as something innate but rather as something acquired in childhood. This is what Bondon (1974) termed as the “primary effect” (see also Jackson and Jonsson 2013). According to this perspective, if we do not include social background variables in the wage equation or treat them as instrumental variables, we will also get biased estimates of the wage premium.
also a rigorous demonstration of how educational expansion shapes the dynamics of changing labor market inequalities.

China is a suitable case for such an intellectual exercise. Over the past three decades, China has had the largest higher education expansion in the world history. Particularly, massive expansion in the late 1990s has generated an over six-fold increase in new college enrollments, from 1.08 million in 1998 to 6.61 million in 2010 (Ministry of Education, 2011; see Table 3.1, Chapter 3). This sharp increase was due to a new educational policy launched in 1999, which aimed to increase the prevalence of college opportunity for the youth population (Ministry of Education 1998). Starting from 2003, i.e. when the first after-expansion college cohort began to work, the labor market also witnessed an increasing supply of college graduates: in 1998, only 3.5 percent of the Chinese workers were college-educated; in 2010, 10.1 percent had received college education (Table 3.5, Chapter 3). Given such dramatic changes in such a short time, the questions become: does the dramatically increased quantity of college graduates equate with a decline in their average ability? How do changes in cohort qualities affect the trend in the college premium? Moreover, how does the college premium vary by different social origin groups? To shed light on these questions, I use the pooled national representative data of the 1996 Life History and Social Change Survey and the 2005-2010 waves of the China General Social Survey (CGSS) to describe the empirical trends and patterns.

One caveat is that China’s educational expansion happened together alongside its market reforms. With the transformation from a redistributive economy to a market economy, the demand for skilled workers increased (Dong and Xu 2009; Liang et al 2009), and the rewarding mechanisms to education have also changed (Bian and Logan 1996; Nee 1989; Nee and Matthews 1996; Xie and Hannum 1996; Xie and Wu 2005; Wu and Xie 2003; Walder 1992,
2002; Zhou 2000a, 200b). As a result, on the one hand, the institutional change of market transition in China tends to advance returns to college education. On the other hand, the increasing supply of skilled workers provided by educational expansion tends to put downward pressure on the college premium. It is still not clear how the price effect might change due to the two competing forces that drive the college premium (Heckman and Yi 2012; Wang 2011).

In the following sections, I will first briefly review the general literature in labor market studies and the literature on China. Then I formulate hypotheses based on the theoretical framework proposed in Chapter 2. Next, I provide descriptive analyses on the trends of earning returns to college education under educational expansion. I will then demonstrate the causal effect of college on earnings and the policy effect of educational expansion on the changing college premium, with a focus on the social origin differentials on the pure premium in recent years.

LITERATURE REVIEW

Skills, Ability, and the Changing College Premium

Earnings return to education is one of the central topics in studies of the labor market (Card 1999; Hout 2012; Mincer 1974). Given that education is concurrent with self-selection (Griliches and Mason 1972; Willis and Rosen 1979), a large body of literature has focused on estimating returns to education, controlling for various forms of selection bias (e.g., Ashenfelter and Rouse 1998; Blundell et al 2005; Card 1999, 2001; Carneiro et al 2010, 2011; Dickson et al 2011; Isacsson 2004; Kaymak 2009; Nordin 2007). A second branch of studies aims to document and explain the trends in wage inequality (e.g., Goldin 2007, 2008; Goldin and Katz 1996, 1998, 1999; Juhn et al 1993; Katz and Murphy 1992), with much thought placed on the demand-side

Recently, special attention has been paid to the supply-side effects. Katz and Murphy’s (1992) study, for example, has argued that the growth rate of the supply of college graduates has played a significant role in explaining the trends in the relative earnings of college graduates in the U.S. from 1963 to 1987. Card and Lemieux (2001) have further shown that the rise in education-related wage differentials is mostly due to changes in the relative earnings of younger college graduates. Using a model with imperfect substitution between workers from different age groups, they have attributed the rising college premium for younger workers to the slowdown in the growth rate of educational attainment starting with the 1950 birth cohorts.

Nevertheless, many of the recent studies are keen on disentangling the sources of the changing college premium. Built on the notion that the unobservable ability can confound the estimation of the college premium (in other words, the “ability bias”), scholars have not yet formed a consensus regarding the relative explanatory powers of the various sources of the changing returns to college. Under a standard wage equation including both education and ability parameters, changes in the college premium can be decomposed into three components: (1) changes in the causal effect of college over time; (2) changes in the return to unobserved ability over time; and (3) changes in the ability differential, conditioning on education outcomes across cohorts (Blackburn and Neumark 1993; Deschênes 2006; Taber 2001).

Researchers who view education to be an indicator of human capital (Becker 1962; Galor and Moav 2004; Guvenen and Kuruscu 2009; Mincer 1974; Rosen 1977) tend to emphasize the first component. They believe that the increase in college premium is due to the improved productivity through education rather than the changing distribution of unobserved ability. For
example, using 1990 U.S. Census data, Fang (2006) finds that college education enhances attendees’ productivity by about 40%, and such productivity enhancement accounts for about two-thirds of the college wage premium. Similarly, Deschênes (2007) reports that after controlling for the effects of unobserved ability and comparative advantage biases across cohorts, the causal return to education increased by approximately 40% in the U.S. between 1979 and 2002. He thus argues that most of the increase in the college premium is due to an increase in the return to schooling, rather than the return to unobserved ability.

In contrast, scholars who consider education to be more of a signaling device (Bedard 2001; Spence 1973; Stiglitz 1975) pay more attention to the second and third sources. Using test scores to measure ability, early work by Blackburn and Neumark (1993) and Murnane et al (1995) have found that much of the increase in the return to education is attributable to an increase in the return to ability. However, Cawley et al (2000; see also Cawley et al 1997, 1999) have indicated that these results are sensitive to identification assumptions. Alternatively, Chay and Lee (2000) have estimated the returns to unobserved ability by decomposing the within- and between-group variances in the wage dispersion. Their results show that there was an approximately 10-25% rise in the unobserved skill premium during the 1980s. Taber (2001) further argues that an increase in the demand for unobserved ability could play a major role in the growing college premium. Particularly, if individuals with higher ability are more likely to attend college, then the increase in the college premium may be due to an increase in the relative demand for high ability workers rather than for college credentials. Based on such evidences of the increasing return to unobserved ability, some scholars have argued that the second source of the effect is important.

However, in both Chay and Lee’s (2000) and Taber’s (2001) studies, one limitation is that
the ability differential between college and high school graduates is assumed to be constant over time. Such limitation has been pointed out by scholars who have argued that the third source is the most significant. Specifically, researchers have found that there are changes in the skill compositions over time: increases in college enrollment lead to changes in the distribution of ability among both college and high school graduates (e.g., Rosenbaum 2003).

Even so, whether this composition effect component is significant is still an open debate. On one hand, some research has argued that the composition effect cannot account for the observed rising college wage premium in the US. According to Blackburn and Neumark (1993, 1995), an increased supply of college graduates is likely to reduce the ability gap between college and high school graduates, and hence depresses the college wage premium, which contradicts the observed pattern. Using the 1940-1990 U.S. census data, Juhn et al (2005) have found that while cohort quality declines as college enrolment increases, there is only a small and marginally significant cohort quality effect on the changes in the college premium.

On the other hand, other recent studies have found more evidence in support of the compositional effect. Contrary to Juhn et al (2005), Rosenbaum’s (2003) results suggest that changes in the ability composition across educational groups can explain almost one half of the increase in the college wage premium during the 1969-1989 period. He has therefore argued that increases in education differentials have been overstated and that the wages of less-educated workers have not “deteriorated” as long as the composition of educational groups is controlled. Carneiro and Lee (2007) have also found the composition effect to be large: in the U.S. of the 1990s, a 14-percent increase in college participation would reduce the college premium by 11 percent. Using a counterfactual framework and controlling for the effect of selection bias, they have found that the price of the unobservable skills and its trend had been underestimated in
previous studies (e.g., Chay and Lee 2000). In their subsequent paper (Carneiro and Lee 2011), they have further shown that the decline in the quality of the college graduates between 1960 and 2000 in the U.S. has led to a decrease of 9 percentage points in the college premium, which was about half of the percentage of growth in college premium in the period. Therefore, rather than providing a single explanation based on the composition effect, their approach attempts to correct the overestimation of the effect of the relative supply change in explaining changes in the college premium for young and old workers. A similar bias-correction analysis is conducted by Kurtzon (2012), in which the composition effect was viewed as a function of “percent educated”. The results have shown that the composition effect could have reduced the measured education premium by 2-20% over the course of the late 20th century.

Furthermore, an additional line of research tries to understand the composition effect through an explicit signaling perspective. Here, the interpretation provided on the shifting college premium is different from the standard supply and demand framework. A recent paper by Zheng (2009) has argued that increased access to college may increase the signaling value of higher education and decrease the signaling value of a high school diploma. Based on the model calibration, it reports that the composition effect accounts for around 17% of the growth in college premium for the 23-26 age cohorts in the US during 1972-2005. This perspective on the declining signaling value of the less-educated group is also shared by Balart (2010). His model suggests that the increase in returns to college is not due to an increase in college graduates’ wage, but rather to a stronger reduction in poorly educated workers’ wage. Last but not the least, also using the signaling framework, Ireland et al (2009) analyzed the composition effect with early-career data. They found that returns by degree class increased across cohorts during a period of substantial graduate expansion in Britain.
Notably, despite the debate on the relative strength of the three sources of the effects of the changing college premium, most of the current empirical studies on the topic are drawn from the experiences of industrial societies, such as the US, UK, and Canada (e.g., Card and Lemieux 2001). International comparison of the US with other OECD countries shows more variation in the wage premium inequality structure (Freeman et al. 2000). Particularly, while changes in the skill demand insignificantly differ among industrial countries, changes in the supply of skilled labor diverge (Acemoglu 2005; Card and Lemieux 2001; Katz and Autor 1999). In addition, nation-specific institutional factors such as the wage setting institution (such as the minimum wage; Addison et al. 2009, 2013; Lee 1999), collective bargaining structure (such as unionization, Kahn 2000), and welfare state regime are also found to affect the trend of educational returns.

**Changing Returns to Education in Reforming China**

There is no comparable case like China in the last decade: the dramatic college expansion has substantially sky-rocketed the supply of college graduates in a relatively short time. China’s specific institutional feature is also well-known: it is under the transition from a redistributive economy to a market economy (Walder 1986; Nee 1989). Since China has experienced both educational expansion and market transition in the reform periods, it is important to conceptually distinguish the two sources of effects on the labor market educational returns.

Most of the existing sociological research looks at the Chinese pattern of changing returns to education from the perspective of market transition. Nee’s market transition theory argues that the development of market institutions strengthens the importance of markets and creates a different incentive structure that better awards education in the market sector (Cao and Nee 2000; Nee and Matthews 1996). Other interpretations based on models of power persistence (Bian and
Logan 1996), power conversion (Stark 1992), the coevolution of state and market (Parish and Michelson 1996; Zhou 2000a), economic growth (Walder 2002) and labor market expansion (Xie and Hannum 1996) argue against the generic effect of market. Wu and Xie (2003) further demonstrate that the sorting process of heterogeneous groups of workers in the labor market, rather than the market per se, explains the higher returns to education in the market sector.

Few sociological studies so far have directly examined the changing labor market returns in the Chinese context through the perspective of educational expansion. To my best knowledge, Hu and Hiebel’s (2013; see also Hu and Hiebel 2014; Zhou 2014) study is the only exception. Based on the urban sample of China General Social Survey (CGSS) 2003 and 2010 waves, their analyses show that the estimated yearly return to education grows from 7.2 percent to 12.6 percent, and the returns to higher education became more positively associated with income in 2010. They then interpret the rising college premium as the result of “the increasing demand for highly skilled labor outstrips the concurrent increasing supply of college-educated workers” (p. 8). However, both the results and the conclusion are misleading. First, their research design can hardly detect the treatment effect of the 1999 educational expansion. The birth cohorts (1971-1978 vs. 1978-1985) they compare are not exactly before- and after-expansion cohorts. Second, results based on the two survey years distort the overall trend of changing college premium. As several recent studies show, returns to college have declined after 2003 (Li et al 2013; Liu and Zhang 2008). There is even evidence that the return to college is lower in 2009 than in 2004 in urban China (Liu and Zhang 2008: Table 2 and Figure 4). Third, given the high unemployment rate among the college graduates (Cai and Wang 2010; Li 2014), a standard supply-demand

49 Tsai and Xie (2008, 2011) have studied the effects of higher education on earnings in the Taiwanese context. However, unlike the sudden shock that is China’s educational expansion in 1999, which has a nature of quasi natural experiment, there is no clear-cut policy effect in the Taiwanese context. Nor is the educational opportunity structure similar in the two contexts.
framework can hardly explain the rising college premium. As I will show in the analyses, labor market sorting and signaling can provide a more plausible story. Finally, the use of the urban sample tends to bias their estimated schooling returns because rural migrants can compete with urban workers in cities (Demurger et al 2012).

The topic of China’s changing returns to education has also attracted substantial interest among economists (e.g., Chen and Li 2014; Chen and Hamori 2009; Heckman and Li 2004; Heckman and Yi 2012; Li et al 2012; Li and Zhang 2010; Meng and Kidd 1997; Knight and Song 2003; Ning 2010; Ren and Miller 2007; Zhao and Zhou 2002; Zhong 2011; Zhu 2011). However, due to the time span of the available data covered, most of the research could only estimate and infer the pre-2002 trends. According to a recent meta-analysis on this topic, only two out of 43 economic papers have extended the time covered to post-2002 years (Liu and Zhang 2008: Table 1). Therefore, we still know little about the trend of the returns to education after the 1999 massive expansion is (from 2003 onwards), and the effect of educational expansion on the college premium is yet to be investigated.

In this chapter, I move towards that direction and study how educational expansion affects changes in the college premium. Considering the effect heterogeneity in the college premium (Brand and Xie 2010; Tsai and Xie 2008, 2011), I investigate how changes in the college premium vary by social origins. I also take this opportunity to evaluate the theoretical debate addressed above and understand the sources and mechanisms of the changing college premium in China. Based on my theoretical framework in Chapter 2, I propose a sorting and signaling story to understand the changing college premium.

---

50 The most widely-used data are the 1988, 1995 and 2002 waves of China Household Income Project (CHIP) data (e.g., Yang 2006; Wang et al 2014). Others include data from the 2000 Urban Household Income and Expenditure Survey (Heckman and Li 2004), the Chinese Twins Survey (Li et al 2012), the China Health and Nutrition Survey (Liu and Zhang 2008), etc.
EDUCATIONAL EXPANSION AND CHANGING COLLEGE PREMIUM IN REFORMING CHINA: HYPOTHESES

Over the past three decades, China had expanded educational opportunities at all levels with varying degrees, resulting in an apparent “bottleneck” structure, as described in Chapter 3. The compulsory education has been provided universally since the mid-1990s, and the college opportunity has been widely available to senior high school graduates since the 1999 educational reform. In contrast, senior high school opportunity is still limited: the transition rate declined from 40.9 percent in 1978 to 22.8 percent in 1998, and increased to 47.8 percent in the first decade of the 21th century (Table 3.1, Chapter 3). As a result, the labor supply has shifted in a dramatic way, particularly for college graduates after the 1999 expansion. Figure 6.1 plots the labor market participation rates (% of the total labor force) for upper secondary school and college graduates since the late 1990s.

It shows that the participation rate for upper secondary school graduates has remained rather stable: even for the 20-24 age groups, the rate only fluctuates around 18 percent. College participation has increased much faster, especially for the young age groups: in 2002, 9.5 percent of the 25-29 age groups had been college-educated; in 2010, the rate jumped to 20.7 percent (see also Table 3.5, Chapter 3). Thus, the relative supply of college graduates has largely increased for the young cohorts.
The sharp increase of college opportunity for the young cohorts has significantly changed the ability composition of college graduates in the last decade. On one hand, more and more students who would have been not qualified for college before the expansion are now able to attend college. On the other hand, as college admissions are still mainly based on the standardized College Entrance Exam, a certain level of differentiation and selection is still guaranteed among the after-expansion cohorts. As already discussed in Chapter 3, each year, and within each province, the provincial level educational authorities announce four cutoff points of entrance exam scores as the minimum requirement for the admission to first-tier, second-tier and third-tier universities (4-year college, or benke) and junior colleges (3-year college, or zhuankan), respectively (Hannum et al 2012). Students whose exam score is below the cutoff-point
Table 6.1: Educational Expansion for Universities and Junior Colleges, National Statistics, 1998-2010

<table>
<thead>
<tr>
<th>Year</th>
<th>Number of Newly-admitted Students (10 thousand)</th>
<th>Number of College</th>
<th>Teacher-student Ratio a</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>University</td>
<td>Junior College</td>
<td>% Junior College</td>
</tr>
<tr>
<td>1998</td>
<td>65.3</td>
<td>43.0</td>
<td>39.7</td>
</tr>
<tr>
<td>1999</td>
<td>93.7</td>
<td>61.2</td>
<td>39.5</td>
</tr>
<tr>
<td>2000</td>
<td>116.0</td>
<td>104.6</td>
<td>47.4</td>
</tr>
<tr>
<td>2001</td>
<td>138.2</td>
<td>130.0</td>
<td>48.0</td>
</tr>
<tr>
<td>2002</td>
<td>158.8</td>
<td>161.7</td>
<td>50.5</td>
</tr>
<tr>
<td>2003</td>
<td>182.5</td>
<td>199.6</td>
<td>52.2</td>
</tr>
<tr>
<td>2004</td>
<td>209.9</td>
<td>237.4</td>
<td>53.1</td>
</tr>
<tr>
<td>2005</td>
<td>236.4</td>
<td>268.1</td>
<td>53.1</td>
</tr>
<tr>
<td>2006</td>
<td>253.1</td>
<td>293.0</td>
<td>53.7</td>
</tr>
<tr>
<td>2007</td>
<td>282.1</td>
<td>283.8</td>
<td>50.2</td>
</tr>
<tr>
<td>2008</td>
<td>297.1</td>
<td>310.6</td>
<td>51.1</td>
</tr>
<tr>
<td>2009</td>
<td>326.1</td>
<td>313.4</td>
<td>49.0</td>
</tr>
<tr>
<td>2010</td>
<td>351.3</td>
<td>310.5</td>
<td>46.9</td>
</tr>
</tbody>
</table>


a The ratio refers to the average number of student a teacher teaches. Calculations are based on currently available statistics.

requirement for second-tier universities can only go to third-tier universities or junior colleges at best, which are generally deemed as not only more expensive but also less prestigious.

According to the first panel in Table 6.1, the number of newly-admittedly students each year increases fast for junior colleges. In 2010, 3.1 million high school graduates went to junior colleges, which accounts for approximately half of the total college admission in that year.

Therefore, although many more senior high school graduates can enter college after the 1999 expansion, many of them are going to the less-prestigious ones, e.g. 3-year junior colleges.  

51 Based on the educational history information in the 1996 Life History and Social Change Survey and the 2008 China General Social Survey, the weighted distribution of college graduates by level of college in 1996 was: 23.5 percent from universities administered by the central state or Ministry of Education, 44.9 percent from universities administered by the provincial government, 6.5 percent from colleges administered by the local government, and 25.1 percent from other universities. In 2008, the percentage becomes 19.2 percent, 40.9 percent, 20.5 percent and 19.3 percent, respectively. The data show that the proportion of students graduating from local colleges increases
Not only have there been changes in students’ ability post-expansion, the quality of college has also declined sharply. According to the second panel in Table 6.1, along the process of college expansion, the total number of 3-year junior college rises quickly, from 1,022 in 1998 to 2,358 in 2010, whereas the number of four-year universities has nearly doubled, from 633 in 2003 to 1,112 in 2010. Since a large proportion of universities had just been rebranded from junior colleges (People’s Daily Online, 30/05/2014), the average college quality declines substantially after the 1999 expansion. The third panel presents the national statistics on the teacher-student ratio, often treated as an indicator of school quality. The college teacher-student ratio has doubled in the twelve years, from 8.38 in 1998 to 16.62, thus pointing to a significant substantially (from 6.5 percent to 20.5 percent). Most of the local colleges are third-tier universities or junior colleges, which have absorb a large number of low-ability students.
decline in the quality of college education. The decline is more serious for junior college than for university. The teacher-student ratio in junior college in 2010 is 23.9, about 1.8 times that in university. Figure 6.2 further graphs the trend of teacher-student ratio in the whole reform period. Clearly, as compared to that of the 1990s, the average quality of senior high school and college has both declined. Starting from 2008, the teacher-student ratio for colleges has been close to that of senior high schools.

The influx of low-ability students into college, especially into junior college, has significantly reduced the social prestige held by college students, which had long been referred as “the ones favored by heaven” before the expansion. Furthermore, the declining college quality further devalues the college credential. When the young cohorts go to the labor market, they face a different labor market opportunity structure from the before-expansion cohorts. Evidence suggests that the unemployment rate among college graduates has escalated (Li et al 2013). According to a report by Chinese Academy of Social Sciences, in 2010 the unemployment rate for college students would be 12 percent, and the cumulative number of unemployed students would be over 13.5 million (Ru et al 2009).

The high unemployment rate among the college graduates is mainly due to unbalanced supply and demand for college graduates (3-year junior college students in particular). Whereas the supply of college graduates increased sharply after the 1999 expansion, the demand for skills also shifted. Table 6.2 presents the percentages of educated workers by industry and by occupation in 2002 and 2010 (see also Murphy and Welch 1993; Katz and Murphy 1992). It shows that except for a few industries, the within-industry relative demand is mostly moderate for junior college graduates. In the meanwhile, the demand for university graduates increases in industries such as IT, banking, business service, and education, and the increase is substantially
Table 6.2: Employment Percentage by Levels of Education and Industry, National Statistics, 2002-2010

<table>
<thead>
<tr>
<th>Sectors</th>
<th>Senior high</th>
<th></th>
<th></th>
<th>3-year Junior college</th>
<th></th>
<th>4-year University a</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Industries</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Agriculture</td>
<td>4.5</td>
<td>5.8</td>
<td>1.3</td>
<td>0.2</td>
<td>0.5</td>
<td>0.3</td>
<td>0.0</td>
</tr>
<tr>
<td>Mining</td>
<td>22.3</td>
<td>23.0</td>
<td>0.7</td>
<td>3.4</td>
<td>8.7</td>
<td>5.3</td>
<td>0.8</td>
</tr>
<tr>
<td>Manufacturing</td>
<td>24.7</td>
<td>20.1</td>
<td>-4.6</td>
<td>4.7</td>
<td>6.4</td>
<td>1.7</td>
<td>1.5</td>
</tr>
<tr>
<td>Electricity and water</td>
<td>39.2</td>
<td>33.1</td>
<td>-6.1</td>
<td>14.3</td>
<td>22.0</td>
<td>7.7</td>
<td>3.9</td>
</tr>
<tr>
<td>Construction</td>
<td>15.0</td>
<td>12.5</td>
<td>-2.5</td>
<td>3.6</td>
<td>3.9</td>
<td>0.3</td>
<td>1.4</td>
</tr>
<tr>
<td>Transportation</td>
<td>29.7</td>
<td>24.1</td>
<td>-5.6</td>
<td>5.6</td>
<td>7.4</td>
<td>1.8</td>
<td>1.4</td>
</tr>
<tr>
<td>IT</td>
<td>35.4</td>
<td>24.3</td>
<td>-11.1</td>
<td>25.6</td>
<td>27.0</td>
<td>1.4</td>
<td>15.9</td>
</tr>
<tr>
<td>Retail and Wholesale</td>
<td>27.4</td>
<td>25.8</td>
<td>-1.6</td>
<td>4.1</td>
<td>8.7</td>
<td>4.6</td>
<td>0.9</td>
</tr>
<tr>
<td>Hotel, Catering</td>
<td>20.8</td>
<td>21.3</td>
<td>0.5</td>
<td>2.5</td>
<td>4.8</td>
<td>2.3</td>
<td>0.5</td>
</tr>
<tr>
<td>Banking</td>
<td>37.1</td>
<td>24.2</td>
<td>-12.9</td>
<td>34.9</td>
<td>32.6</td>
<td>-2.3</td>
<td>10.7</td>
</tr>
<tr>
<td>Real Estate</td>
<td>36.0</td>
<td>27.4</td>
<td>-8.6</td>
<td>23.2</td>
<td>18.3</td>
<td>-4.9</td>
<td>9.1</td>
</tr>
<tr>
<td>Lease, Business service</td>
<td>30.2</td>
<td>24.3</td>
<td>-5.9</td>
<td>8.8</td>
<td>20.7</td>
<td>11.9</td>
<td>2.3</td>
</tr>
<tr>
<td>Research</td>
<td>30.0</td>
<td>18.7</td>
<td>-11.3</td>
<td>24.9</td>
<td>23.9</td>
<td>-1.0</td>
<td>21.7</td>
</tr>
<tr>
<td>Environmental Management</td>
<td>32.5</td>
<td>22.2</td>
<td>-10.3</td>
<td>16.3</td>
<td>13.9</td>
<td>-2.4</td>
<td>5.7</td>
</tr>
<tr>
<td>Public service</td>
<td>20.6</td>
<td>20.4</td>
<td>-0.2</td>
<td>3.4</td>
<td>4.4</td>
<td>1.0</td>
<td>0.7</td>
</tr>
<tr>
<td>Education</td>
<td>33.0</td>
<td>18.0</td>
<td>-15</td>
<td>39.7</td>
<td>33.3</td>
<td>-6.4</td>
<td>17.5</td>
</tr>
<tr>
<td>Health, Security, Welfare</td>
<td>39.9</td>
<td>27.5</td>
<td>-12.4</td>
<td>24.8</td>
<td>33.9</td>
<td>9.1</td>
<td>10.3</td>
</tr>
<tr>
<td>Culture and Entertainment</td>
<td>36.9</td>
<td>25.2</td>
<td>-11.7</td>
<td>26.4</td>
<td>18.6</td>
<td>-7.8</td>
<td>10.2</td>
</tr>
<tr>
<td>Government, Institutions</td>
<td>33.5</td>
<td>23.0</td>
<td>-10.5</td>
<td>37.3</td>
<td>31.6</td>
<td>-5.7</td>
<td>13.8</td>
</tr>
<tr>
<td>International Organizations</td>
<td>16.3</td>
<td>15.3</td>
<td>-1.0</td>
<td>13.7</td>
<td>15.0</td>
<td>1.3</td>
<td>11.1</td>
</tr>
<tr>
<td>Occupations</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Head, Manager</td>
<td>29.9</td>
<td>26.0</td>
<td>-3.9</td>
<td>27.2</td>
<td>21.1</td>
<td>-6.1</td>
<td>12.5</td>
</tr>
<tr>
<td>Professionals</td>
<td>31.8</td>
<td>22.6</td>
<td>-9.2</td>
<td>24.0</td>
<td>31.3</td>
<td>7.3</td>
<td>12.5</td>
</tr>
<tr>
<td>Clerks</td>
<td>37.8</td>
<td>25.6</td>
<td>-12.2</td>
<td>23.9</td>
<td>26.1</td>
<td>2.2</td>
<td>6.7</td>
</tr>
<tr>
<td>Business service</td>
<td>23.4</td>
<td>24.3</td>
<td>0.9</td>
<td>2.8</td>
<td>7.6</td>
<td>4.8</td>
<td>0.4</td>
</tr>
<tr>
<td>Agriculture, Irrigation</td>
<td>4.4</td>
<td>5.8</td>
<td>1.4</td>
<td>0.1</td>
<td>0.5</td>
<td>0.4</td>
<td>0.0</td>
</tr>
<tr>
<td>Production, Transport</td>
<td>22.1</td>
<td>17.9</td>
<td>-4.2</td>
<td>1.9</td>
<td>3.8</td>
<td>1.9</td>
<td>0.2</td>
</tr>
<tr>
<td>Others</td>
<td>17.7</td>
<td>20.5</td>
<td>2.8</td>
<td>3.2</td>
<td>7.4</td>
<td>4.2</td>
<td>0.9</td>
</tr>
</tbody>
</table>


Note: National statistics for between-industry share is not available in 2010.

a Not including graduates.
larger than that for junior college graduates. Comparing the demand shift for high school graduates and university graduates, we find that the increasing demand for university students correspond with the declining demand for senior high school graduates, which suggests potential skill replacement (Autor and Katz 1998).

Similarly, the demand shifts by occupation also show that the demand for university graduates increases much faster than for junior college graduates, and it is biased toward manager, professional, and clerk positions. Given the different paces of growth in the within-industry and within-sector demand for skills, it is reasonable to speculate that the 3-year junior college graduates and 4-year university students have different rates of educational returns, which is supported by recent studies on the effects of the quality of Chinese colleges (Li 2003; Li et al 2012; Zhong 2011). Moreover, as the labor demand for workers with a junior college credential increases slowly, a certain group of junior college students tends to be over-educated (for a review of the over-education literature, see Groot and Brink 2000; Walker and Zhu 2005).

Since the after-expansion cohorts are characterized by a larger quantity and a lower quality on average, the price for college education tends to decline (quantity), and the market value signaled by the college credential also tends to be lower (quality, or unobserved ability). If this is the case, we would observe:

**Hypothesis 1: College premium becomes lower for the after-expansion cohorts.**

In the hypothesis, the after-expansion cohorts refer to those who finish senior high school education in 1999 or afterwards. We also may expect that the within-group residual, or the income returns to unobserved ability, becomes smaller for the after-expansion cohorts.

While **Hypothesis 1** predicts that educational expansion negatively impacts the college premium, such negative effects hold only as aggregate effects. Because income returns to
education are also shaped by individual level factors such as ability and resources, the general negative impact educational expansion brings is compensated differently for various social groups. If it is the case that the higher class gets disproportionally larger advantages not only in educational attainment (Chapter 4) but also in job placement (Chapter 5), the linkage between college education and destination is stronger for the higher class than the lower class. Provided that desirable positions are often associated with good income, it is reasonable to predict that:

_Hypothesis 2: After the expansion, the decline in the college premium is less salient for college graduates with advantaged social origin._

Nonetheless, the return to education presents not simply the premium of educational credentials rewarded in the labor market, but also the effect of individuals’ unobserved heterogeneity, say, ability differentials (Card 1999; Rosenbaum 2003). The composition effect could bias our estimation through two mechanisms.

First, the signaling of college credentials may work differently for people from various social groups. As Chapter 4 shows, China’s “bottleneck” structure has disproportionally helped low-ability children with advantageous background to go to college. Given the widely available college opportunity for children from better-off families, having a college degree may not signal high ability, but not having college education would be a clear sign of low ability. In the opposite, for children with disadvantaged backgrounds, the college credential would be a sign for high ability, and not having a degree would not signal low-ability. If this is the case, then we may expect the decline in the college premium differs by social origin groups.

Second, as Brand and Xie (2010) indicate, individuals who are least likely to obtain a college credential have a larger treatment effect once they have the college credential. While under recent educational expansion there is a relatively larger probability of children from the
higher class than from the lower class to enter college, and a relatively higher propensity of children from the higher class to use college education for desirable jobs, it is still unclear about the direction of selectivity for different classes and the general patterns of origin-based college premiums across time. One possibility is that because the two processes are more selective for children from the lower class, unobserved factors such as ability matter more for them. If this is true, then a pattern of negative selection is plausible: going to college per se benefits children from the lower class more once they have successfully overcome the barriers in educational and occupational attainment, but the benefit is lower after the expansion.

DATA, VARIABLES, AND METHODS

Data

My first aim is to show the trends of earning returns to education across time. Despite the existence of a multitude of research on schooling returns in China (Liu and Zhang 2008), no study to date has utilized multiple datasets to depict the full-ranged trends of schooling returns in the reform periods. To get a whole picture of such trends in the last three decades, I employ a variety of widely-used survey data to estimate schooling returns in a particular survey year or a retrospective year.\footnote{The selected data include the 1994 \textit{State and Life Chances in Urban China} survey (Zhou 1998), the 1996 \textit{Life History and Social Change} survey (Treiman and Walder 1998), the 2000 \textit{Life Changes in Urban China} survey (Li 2001, 2002), the 1988, 1995, and 2002 waves of the \textit{Chinese Household and Income Project} (Griffin and Zhao 1993), the 1991 to 2009 waves of the \textit{China Nutrition and Health Survey} (Liu and Zhang 2008), and the 2003 to 2010 waves of the \textit{China General Social Survey} (Wang 2011). See the Appendix for more details.} Table 6.1 in the Appendix gives a brief summary of the sampling population of each dataset and the years of income information available. To be consistent across datasets, analyses are limited to the urban working population aged 20-60.

My second aim is to illustrate the relationship between educational expansion and the changing college premium in the reform periods. For this part of the analysis, I use the pooled
national sample of the 1996 *Life History and Social Change Survey* and the 2005-2010 waves of the *China General Social Survey*. All the datasets are nationally representative and highly-comparable. A detailed description of these datasets can be found in Treiman and Walder (1998) and Wang (2011). To investigate the pattern of changing college premium under educational expansion, I restrict the sample to senior high school and college graduates (not including graduate school attendees) who are at the risk of college transition during 1981-2010. Considering that the regular age of making college transition is 18 in China, I exclude the early (before 1963) and late birth cohorts (survey year-birth<18) from each dataset. These procedures have generated a sample of 3,236 senior high school graduates and 2,789 college graduates.

**Variables**

For all the survey data, the dependent variable is the logged monthly income in a specific year, which is adjusted by the 1996 consumption price index (CPI).\(^{53}\) In the first part of analyses on the trend of schooling returns, to ensure the results are comparable across datasets, only a minimum set of covariates are included in a basic Mincer equation. The independent variables include sex (male=1), age, age-squared, education and party membership, with provincial dummies as controls.\(^{54}\) Party membership is included to reflect the aspect of human capital associated with political advantage (Xie and Hannum 1996). It is a dummy variable (yes=1). Education is measured both in years of schooling and in levels. When studying income gaps

\(^{53}\) Information on wage and work hours is not available in certain dataset (e.g., CGSS 2003, 2005), so I use monthly income to ensure the consistency across datasets. Moreover, monthly income is a better indicator than hourly income in the Chinese context because of the additional unreliability generated by the self-reported working hours.

\(^{54}\) Work experience is better than age for a basic Mincer equation. However, it is based on the report of first job entry year, which is not available in multiple datasets (e.g., State 1994, CHIP 1988 and 1995 waves, and CHNS datasets). We can calculate it as age minus years of schooling minus six (Tsai and Xie 2008), but by doing so the reliability of this measure decreases, particularly in retrospective years.
between high school and college graduates, senior high school is treated as the reference group and the distinction between 3-year junior college and 4-year university is made (0=senior high school, 1=junior college, 2=university).

In the second part of the analyses, college premium is strictly defined as the income differential between full-time college and senior high school graduates.\(^{55}\) Therefore, in a causal inference framework, the assignment takes place at the time of college transition. Those who go to full-time college belong to the treatment group, and those who do not make successful transition to full-time college go to the control group. The control group also includes people who pursue adult college or part-time college (not in graduate school) after work, as they do not manage to enter college at the year of college transition.\(^{56}\) For each respondent, the year of college transition is obtained from the reported educational history or the enrollment status and specific grade if educational history information is not available. I divide the year of college transition into three periods: 1981-1992, 1993-1998, and 1999-2010. The year 1992 witnessed Deng Xiaoping’s political tour to southern China, after which China deepened its economic reforms and marketization. It also reflects the start of changes in college tuition and in college employment (Li et al 2012, 2013). The year 1999 marked the implement of an educational policy designed for massive college expansion. Comparing the college premium in the second and third

\(^{55}\) Other ways to measure the college premium could be the income gap between college and non-college graduates (e.g., Tsai and Xie 2008, 2011), or between 4-year university and senior high school graduates (Wang et al 2014). The first measure is misleading because educational attainment is a sequence of schooling transitions (Mare 1980, 1981). Those who do not finish senior high school would never enter the risk set of making college transitions. The second measure neglects the fact that there are different types of colleges, such as the 3-year junior college in China and the two-year college in the US.

\(^{56}\) The process of college attendance for these people is different from that for regular full-time college students. Excluding these cases from the analyses does not significantly affect the empirical patterns. Results are available upon request.
Table 6.3: Descriptive Statistics, by Period of College Transition, the Pooled Sample, 1981-2010

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total</td>
<td>College</td>
<td>HS</td>
<td>Total</td>
<td>College</td>
<td>HS</td>
<td>Total</td>
<td>College</td>
</tr>
<tr>
<td>N</td>
<td>2,797</td>
<td>772</td>
<td>2,025</td>
<td>1,387</td>
<td>620</td>
<td>717</td>
<td>1,891</td>
<td>1,397</td>
</tr>
<tr>
<td>%</td>
<td>46.4</td>
<td>25.0</td>
<td>75.1</td>
<td>22.2</td>
<td>43.6</td>
<td>56.4</td>
<td>31.4</td>
<td>70.7</td>
</tr>
<tr>
<td>Male %</td>
<td>55.6</td>
<td>58.4</td>
<td>54.6</td>
<td>53.0</td>
<td>58.3</td>
<td>49.0</td>
<td>50.0</td>
<td>49.7</td>
</tr>
<tr>
<td>Party %</td>
<td>22.8</td>
<td>33.8</td>
<td>19.2</td>
<td>16.4</td>
<td>24.2</td>
<td>10.4</td>
<td>13.4</td>
<td>16.5</td>
</tr>
<tr>
<td>Hukou origin %</td>
<td>64.2</td>
<td>73.8</td>
<td>60.4</td>
<td>67.6</td>
<td>75.9</td>
<td>60.3</td>
<td>72.7</td>
<td>77.7</td>
</tr>
<tr>
<td>Age</td>
<td>38.0</td>
<td>37.3</td>
<td>38.3</td>
<td>30.3</td>
<td>30.1</td>
<td>30.4</td>
<td>24.5</td>
<td>24.4</td>
</tr>
<tr>
<td></td>
<td>(4.87)</td>
<td>(4.93)</td>
<td>(4.82)</td>
<td>(3.62)</td>
<td>(3.30)</td>
<td>(3.95)</td>
<td>(3.47)</td>
<td>(3.65)</td>
</tr>
<tr>
<td>Experience</td>
<td>16.74</td>
<td>15.3</td>
<td>17.28</td>
<td>8.59</td>
<td>8.03</td>
<td>9.09</td>
<td>3.62</td>
<td>3.22</td>
</tr>
<tr>
<td></td>
<td>(5.94)</td>
<td>(5.43)</td>
<td>(6.04)</td>
<td>(4.05)</td>
<td>(4.17)</td>
<td>(3.87)</td>
<td>(3.59)</td>
<td>(3.68)</td>
</tr>
<tr>
<td>Income (logged)</td>
<td>7.13</td>
<td>7.52</td>
<td>6.97</td>
<td>7.37</td>
<td>7.69</td>
<td>7.10</td>
<td>7.22</td>
<td>7.42</td>
</tr>
<tr>
<td></td>
<td>(0.98)</td>
<td>(0.97)</td>
<td>(0.94)</td>
<td>(0.95)</td>
<td>(0.92)</td>
<td>(0.89)</td>
<td>(0.92)</td>
<td>(0.93)</td>
</tr>
<tr>
<td>Father’s education</td>
<td>7.71</td>
<td>9.07</td>
<td>7.19</td>
<td>9.00</td>
<td>9.74</td>
<td>8.35</td>
<td>10.02</td>
<td>10.40</td>
</tr>
<tr>
<td></td>
<td>(4.28)</td>
<td>(4.49)</td>
<td>(4.08)</td>
<td>(3.56)</td>
<td>(3.53)</td>
<td>(3.46)</td>
<td>(3.22)</td>
<td>(3.21)</td>
</tr>
<tr>
<td>Mother’s education</td>
<td>5.53</td>
<td>7.13</td>
<td>4.92</td>
<td>7.10</td>
<td>7.84</td>
<td>6.47</td>
<td>8.79</td>
<td>9.17</td>
</tr>
<tr>
<td></td>
<td>(4.37)</td>
<td>(4.75)</td>
<td>(4.06)</td>
<td>(4.03)</td>
<td>(4.09)</td>
<td>(3.87)</td>
<td>(3.70)</td>
<td>(3.63)</td>
</tr>
<tr>
<td>Father’s ISEI</td>
<td>39.7</td>
<td>44.5</td>
<td>37.9</td>
<td>39.2</td>
<td>42.5</td>
<td>36.4</td>
<td>38.8</td>
<td>40.9</td>
</tr>
<tr>
<td></td>
<td>(17.9)</td>
<td>(18.7)</td>
<td>(17.3)</td>
<td>(17.4)</td>
<td>(18.0)</td>
<td>(16.3)</td>
<td>(16.1)</td>
<td>(16.5)</td>
</tr>
<tr>
<td>Mother’s ISEI</td>
<td>33.2</td>
<td>39.0</td>
<td>31.0</td>
<td>34.0</td>
<td>38.1</td>
<td>31.2</td>
<td>36.8</td>
<td>38.3</td>
</tr>
<tr>
<td></td>
<td>(15.5)</td>
<td>(18.8)</td>
<td>(13.4)</td>
<td>(15.7)</td>
<td>(17.7)</td>
<td>(12.8)</td>
<td>(15.7)</td>
<td>(16.7)</td>
</tr>
<tr>
<td>Number of siblings</td>
<td>1.26</td>
<td>1.11</td>
<td>1.32</td>
<td>0.71</td>
<td>0.55</td>
<td>0.84</td>
<td>0.57</td>
<td>0.53</td>
</tr>
<tr>
<td></td>
<td>(1.72)</td>
<td>(1.52)</td>
<td>(1.79)</td>
<td>(1.17)</td>
<td>(1.01)</td>
<td>(1.28)</td>
<td>(0.95)</td>
<td>(0.91)</td>
</tr>
<tr>
<td>Region %</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>West</td>
<td>22.9</td>
<td>19.5</td>
<td>24.0</td>
<td>24.5</td>
<td>23.5</td>
<td>25.3</td>
<td>23.6</td>
<td>20.4</td>
</tr>
<tr>
<td>Central</td>
<td>30.8</td>
<td>23.6</td>
<td>33.2</td>
<td>23.1</td>
<td>17.7</td>
<td>27.2</td>
<td>20.8</td>
<td>21.9</td>
</tr>
<tr>
<td>East</td>
<td>46.4</td>
<td>57.0</td>
<td>42.8</td>
<td>52.4</td>
<td>58.8</td>
<td>47.5</td>
<td>55.6</td>
<td>57.7</td>
</tr>
</tbody>
</table>

Note: Reported are weighted sample percentages, means, and standard deviations.

*a Standard deviations in parentheses.

time periods thus provides evidence for the policy effect. In the analyses, I look at the college
premium both for the 3-year junior college graduates and 4-year university graduates.

Additional independent variables include the respondent’s work experience, current
occupation, and type of work unit (danwei). Family background indicators include the
respondent’s hukou origin, father’s and mother’s years of schooling, father’s and mother’s
occupation when the respondent was aged 14, and number of siblings. Experience is measured
by years of employment since the first job entry year. Occupation is measured by a continuous
ISEI score (Ganzeboom et al 1992). Work unit is a four-category variable: 1=state agencies and
institutions, 2=state-owned enterprises, 3=private enterprises including foreign and join venture
firms, 4=collective-owned enterprises and farm workers. It captures the difference in the
rewarding mechanisms in various sectors (Xie and Wu 2005; Guo and Wu 2009). Hukou origin
is a dummy variable, indicating whether the respondent’s father’s hukou was urban when the
respondent aged 14 (1=urban, 0=rural). To reflect additional regional inequalities in China (Li
and Gibson 2013; Zhang and Zhou 2012), a region variable is constructed based on the official
definition (1=western region, 2=central region, and 3=eastern region). Finally, birth year is
another control variable for cohort heterogeneity (Heckman and Li 2004).

Table 6.3 presents summary statistics of the variables use in this study by periods of college
transition. It shows that there are significant differences not only among the college and senior
high school graduates, but also between the pre- and post-expansion college entry cohorts. The
share of college graduates has increased substantially across periods. In the second period,
college graduates account for 43.6 percent of the sample; after the 1999 expansion, the share
becomes 70.7 percent. Compared to the second period, there are lower percentages of male and
party members among the college graduates in the third period, but the share of people with
urban hukou origin increases. The mean age and experience are declining across college entry
cohorts, as those who have experienced college transition after 1999 are still in their early careers.
In all the periods, the average income of college graduates is significantly higher than that for
high school graduates, but the difference in the mean income becomes lower in the third period.
Parental years of schooling and ISEI scores are also higher for college graduates than for high
school graduates. Compared to the second college entry cohorts, the mean parental education is
larger for the youngest cohorts. Finally, the average number of siblings is lower among the college graduates, and the share of college-educated workers is larger in the Eastern region.

**Methods**

I use a counterfactual causal inference framework to study the causal effect of college on earnings (Dehejia and Wahba 1999; Harding et al. 2007; Gangl 2010; Morgan and Winship 2007). To understand changes in the college premium across college transition cohorts, I further adopt a “difference-in-difference” perspective to detect the policy effect of the 1999 educational expansion. Considering that the expansion policy does not directly affect the entire population but only a certain group of individuals who respond to the increased college opportunity (Carneiro et al. 2010; Heckman and Vytiacil 2001, 2004), I employ a ranking and sorting method to identify the “compliers” under a LATE (local treatment effect) framework (Angrist et al. 1996; Imbens and Agresti 1994; Imbens 2010).

**Identifying the Causal Effect of College on Earnings**

In the counterfactual causal inference literature, different members of a population can have different treatment effects (Heckman and Vytiacil 2005, 2006a; Imbens 2010; Winship and Morgan 2007, 2015). The ATE (average treatment effect) represents the average treatment effect over the entire population. The ATT (average treatment effect on the treated) and ATUT (average treatment effect on the untreated) refer to the average treatment effect for those who are allocated into the treatment group and the control group, respectively. The LATE (local average treatment effect) averages the distribution of treatment effects for those who switch into treatment as a result of a change of the value of some instrument affecting decisions to participate. The MTE (marginal treatment effect) unifies the above evaluation parameters.
(Heckman and Vytiacil 1999, 2004, 2006a), and refers to the effect of a treatment on the marginal individual entering the treatment.

Following the marginal treatment effect framework, consider a model with heterogeneous returns to education for the $i^{th}$ individual:

$$\ln Y_i = \beta_i S_i + \gamma X_i + U_i$$  \hspace{1cm} (1)

where $\ln Y_i$ is logged income, $S_i$ is schooling level, $X_i$ is a vector of independent variables, $U_i$ is the residual term, $\beta_i$ is the rate of return to education, and $\gamma$ is a vector of coefficients. The schooling choice $S_i$ is a dummy variable, with $S_i=1$ denoting college graduates, and $S_i=0$ denoting senior high school graduates. The two potential selection outcomes ($\ln Y_{0i}$ and $\ln Y_{1i}$) can therefore be written as:

\[
\begin{cases}
\ln Y_{0i} = \gamma_0 X_i + U_{0i} & \text{if } S_i = 0 \\
\ln Y_{1i} = \gamma_1 X_i + U_{1i} & \text{if } S_i = 1
\end{cases}
\]  \hspace{1cm} (2)

where $E(U_{0i}|X_i)=0$ and $E(U_{1i}|X_i)=0$ in the population. Based on the potential outcome equations, the observed logged income in equation (1) can be expressed as $\ln Y_i = S_i \ln Y_{1i} + (1 - S_i) \ln Y_{0i}$ and $U_i = S_i U_{1i} + (1 - S_i) U_{0i}$.

Since each individual can choose only one of the above two states, we cannot observe both $\ln Y_{1i}$ and $\ln Y_{0i}$ due to a fundamental missing data problem (Holland 1986). Suppose the schooling choice is determined by the following rule of treatment allocation:

$$S_i^* = p_i Z_i - U_{si}$$  \hspace{1cm} (3)

$$S_i = 1 \text{ if } S_i^* > 0; S_i = 0 \text{ if } S_i^* \leq 0.$$

Here $S_i^*$ is the latent tendency to attend college, $Z_i$ is an observed vector of variables that predict the treatment probability, $p_i$ is a vector of coefficients so that $P_i=p_i Z_i$ denotes the propensity score (Rubin 1974, 1997; Rosenbaum and Rubin 1983) or probability of receiving treatment, which can be estimated by a logit or probit model, and $U_{si}$ is the unobserved heterogeneity for
the $i^{th}$ individual. The school choice decision for individual $i$ is thus determined completely by the comparison of the “propensity score” of receiving college education due to $Z$ variables, $P_i$, with the unobserved resistance to college education $U_{si}$. The higher the propensity score $P_i$ because of $Z$, or the smaller the unobserved resistance to college education $U_{si}$, the more likely the person goes to college. If $p_iZ_i=U_{si}$, then the person is viewed as the “marginal” individual because whether or not to go to college is indifferent for this person.

Representing the treatment effects for these marginal individuals, the marginal treatment effect (MTE) conditional on $X_i$ and $U_{si}$ is defined as:

$$\text{MTE}(X_i = x, U_{si} = u_s) = \mathbb{E}((\Delta_i | X_i = x, U_{si} = u_s) = \mathbb{E}(\beta_i | X_i = x, U_{si} = u_s) = (\gamma_1 - \gamma_0) x + E(U_{1i} - U_{0i} | U_{si} = u_s)$$

(4)

where $\Delta_i = \ln Y_{1i} - \ln Y_{0i}$ denoting the treatment effect that moves individual $i$ from $S_i = 0$ to $S_i = 1$. If the treatment effect is homogeneous with respect to the unobserved heterogeneity $U_{si}$, a flat line of MTE would be observed as a function of $U_{si}$. Otherwise, whether MTE decreases or increases with $U_{si}$ shows how the latent tendency of college attendance is positively or negatively associated with the treatment effect.

There are several approaches to estimate MTE, including parametric, semi-parametric and non-parametric methods. Heckman and Vytiacil (2006a, 2006b) give an informative instruction on how to estimate MTE under a series of assumptions. Once MTE is known, all the treatment parameters can be obtained using weighted averages of MTE (Heckman and Vytiacil, 2001). A set of weight formulas for the OLS and IV estimators as well as the ATT, ATE, ATUT, and LATE parameters are given in Heckman and Vytiacil (1999, 2000, 2001, 2006b).

Moreover, given the heterogeneous treatment effect captured by MTE, the conventional OLS estimation is biased. Based on equations (2), the probability limit of the OLS estimator is:
\[ plim(\hat{\beta}_{ols}) = E(\ln Y_i | X_i, S_i = 1) - E(\ln Y_i | X_i, S_i = 0) = E(\gamma_1 X_i + U_{1i} | X_i, S_i = 1) - E(\gamma_0 X_i + U_{0i} | X_i, S_i = 0) = \hat{\beta}(X) + [E(U_{1i} | S_i = 1) - E(U_{0i} | S_i = 0)] = ATE + \text{Bias} \] (5)

where ATE is the average treatment effect defined as:

\[ ATE = E(\Delta_i | X_i) = E(\beta_i | X_i) = \hat{\beta}(X). \]

The equation (5) can also be decomposed by ATT so that:

\[ plim(\hat{\beta}_{ols}) = E(\ln Y_i | X_i, S_i = 1) - E(\ln Y_i | X_i, S_i = 0) = E(\gamma_1 X_i + U_{1i} | X_i, S_i = 1) - E(\gamma_0 X_i + U_{0i} | X_i, S_i = 0) = E((\beta_i | X_i, S_i = 1) + [E(U_{0i} | S_i = 1) - E(U_{0i} | S_i = 0)] = \text{ATT} + \text{Selection Bias} \] (6)

where ATT is treatment effect on the treated defined as:

\[ \text{ATT} = E(\Delta_i | X_i, S_i = 1) = E(\beta_i | X_i, S_i = 1) = \hat{\beta}(X) + E(U_{1i} - U_{0i} | S_i = 1) = \text{ATE + Sorting Gain} \] (7)

Based on the decomposition in (5), (6), and (7), the total bias from the OLS estimator would be:

\[ \text{Total Bias} = \text{Selection Bias} + \text{Sorting Gain} = (\hat{\beta}_{ols} - \text{ATT}) + (\text{ATT-ATE}) \] (8)

Here, the selection bias \( E(U_{0i} | S_i = 1) - E(U_{0i} | S_i = 0) \) is the mean difference in the unobserved heterogeneity between the counterfactual of what a college graduate would have earned if he doesn’t attend college and what an average high school graduate earns. The sorting effect \( E(U_{1i} - U_{0i} | S_i = 1) \) is the mean gain of the unobserved heterogeneity for college graduates, i.e. the difference between what an average college graduate earns and the counterfactual of what he would have earned if he doesn’t go to college. The bias in (8) is the sum of selection and sorting bias. In the empirical analyses, I will estimate the treatment effect parameters and decompose the total bias for each period of college transition.
Identifying the Policy Effect on the College Premium

Considering the double-treatment nature of this study, I take a difference-in-difference approach to examine how educational expansion affects the college premium. The key notion is that the expansion policy only moves a certain group of individuals into college. Borrowing the notation from the LATE framework, these individuals would be the compliers who are directly affected by the expansion policy. There are also always-takers who would go to college in the absence of educational expansion and never-takers who would not be able to attend college regardless whether there is expansion. No defiers are considered because under the expansion policy the monotonicity assumption (Hirano et al 2000; Imbens and Agresti 1994) is plausible: no individuals would choose a lower level of education after the expansion than their real (or counterfactual) pre-expansion educational levels.

To identify the always-takers, compliers and never-takers, I adopt a ranking-and-sorting method. I assume that for each college transition cohort individuals can be ranked by their probability (or propensity score) of transitioning to college, and only those whose percentile is above a certain threshold would be selected into college. Similar to equation (2), the selection rule for college in the j^{th} period in the treatment or control group is (for notational simplicity, ignore the subscript of j for the i^{th} person):

\[ S_i^* = \text{Rank}_{p_i z_i} - t_j \]  (9)

\[ S_i = 1 \text{ if } S_i^* > 0; S_i = 0 \text{ if } S_i^* \leq 0. \]
Table 6.4: Construction of the Always-taker, Complier and Never-takers

<table>
<thead>
<tr>
<th></th>
<th>Pre-expansion cohorts</th>
<th>Post-expansion cohorts</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Observed</td>
<td>Counterfactual</td>
</tr>
<tr>
<td>Share of college attendees</td>
<td>$T_j$</td>
<td>$T_{j+1}$</td>
</tr>
<tr>
<td>Always-takers</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Compliers</td>
<td>0 1: The top of observed high school graduates</td>
<td>0 0: The bottom of observed college graduates</td>
</tr>
<tr>
<td>Never-takers</td>
<td>0 0</td>
<td>0 0</td>
</tr>
</tbody>
</table>

Note: 1=college, 0=high school. The counterfactual framework distinguishes compliers from never-takers for the pre-expansion cohorts and always-takers from compliers for the post-expansion cohorts.
where \( \text{Rank}_{p_iZ_i} \) is the percentile of individual \( i \)'s probability (given by the propensity score \( p_iZ_i \)) of college attendance, \( t_j \) is the threshold defined by the level of selectivity during the college transition in the \( j^{th} \) period \((0 < t_j \leq 100)\). For example, for a given period, if we observe that \( t_j \) percent of the total transition cohorts are in college in the \( j^{th} \) period, then only those whose percentile is above \((100 - t_j)\) would be able to successfully enter college.

Let \( T_j \) and \( T_{j+1} \) denote the observed percentage of college attendees in the college transition pool before and after the expansion, respectively, and \( T_j < T_{j+1} \). Now let’s consider two counterfactuals. First, who would be the always-takers, compliers and never-takers for the pre-expansion cohorts if there would be educational expansion? With the assumption that the pre-and post-expansion cohorts are balanced in the absence of the expansion policy, the percentage of college attendees would increase from \( T_j \) to \( T_{j+1} \). This is equivalent to selecting a certain proportion of high school graduates into college based on their probability rank in the control group \((S=0)\).\(^57\) Using a little algebra, the percentile would be \((1 - \frac{T_{j+1} - T_j}{100 - T_j}) \times 100\), and only those high school graduates whose within-group percentile is above it would be selected into college. They would be the compliers, and the remaining senior high school graduates would be the never-takers. The always-takers would be those previous college attendees.

In the second counterfactual case, I ask: who would be the always-takers, compliers and never-takers for the post-expansion cohorts if there would be no educational expansion? This is equivalent to dropping a certain proportion of post-expansion college attendees out of college. Using the within-group probability rank of the treatment group, those whose percentile is lower

\(^{57}\) This approach rests on the assumption that individuals with the same rank, i.e., those in the same place in their cohort's probability of college attendance have about the same level of ability.
Table 6.5: A Difference-in-Different Framework to Interpret the Effects of Educational Expansion on the Changing College Premium

<table>
<thead>
<tr>
<th>Groups</th>
<th>Ability</th>
<th>Pre-Expansion Premium $(D_1)$</th>
<th>Post-Expansion Premium $(D_2)$</th>
<th>Difference-in-difference $(DID); D_2 - D_1$</th>
<th>Difference in difference-in-difference $(DID_1)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Always-takers (A)</td>
<td>High</td>
<td>A vs. N $(D_{1AN})$</td>
<td>A vs. N $(D_{2AN})$</td>
<td>$D_{1AN} vs. D_{2AN}$ $(DID_1)$</td>
<td>$D_{1AN}$ vs. $(DID_1)$ vs. $(DID_2)$ Ability Bias High</td>
</tr>
<tr>
<td></td>
<td></td>
<td>College$<em>1$+ Ability$</em>{High-Low}$</td>
<td>College$<em>2$+ Ability$</em>{High-Low}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Compliers (C)</td>
<td>Median</td>
<td>C vs. N $(D_{1CN})$</td>
<td>C vs. N $(D_{2CN})$</td>
<td>$D_{1CN} vs. D_{2CN}$ $(DID_2)$</td>
<td>$D_{2CN}$ vs. DID$_2$ Ability Bias Median</td>
</tr>
<tr>
<td></td>
<td></td>
<td>College$<em>2$+ Ability$</em>{Median-Low}$</td>
<td>College$_2$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never-takers (N)</td>
<td>Low</td>
<td>DID$_1$ vs. DID$_2$ College$_1$</td>
<td>DID$_1$ vs. DID$_2$ College$_1$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The above identification strategy of the always-takers, compliers and never-takers is summarized in Table 6.4 and intuitively depicted in Figure 6.3. After identifying the three groups, I adopt a difference-in-difference approach to identify the “pure” college premium for the pre- and post-expansion cohorts. With further decomposition of the effects, I am also able to obtain the ability bias in the college premium for the always-taker and compliers. Table 6.5 presents the framework.

To start, always-takers, compliers and never-takers refer to three groups with varying unobserved ability levels (high, median and low, respectively). The college premium observed in pre-expansion period for the always-takers $(D_{1AN})$ is a composite of the pre-expansion treatment before the expansion.

The above identification strategy of the always-takers, compliers and never-takers is summarized in Table 6.4 and intuitively depicted in Figure 6.3. After identifying the three groups, I adopt a difference-in-difference approach to identify the “pure” college premium for the pre- and post-expansion cohorts. With further decomposition of the effects, I am also able to obtain the ability bias in the college premium for the always-taker and compliers. Table 6.5 presents the framework.

To start, always-takers, compliers and never-takers refer to three groups with varying unobserved ability levels (high, median and low, respectively). The college premium observed in pre-expansion period for the always-takers $(D_{1AN})$ is a composite of the pre-expansion treatment before the expansion.

---

58 Since the before-expansion cohorts are more experienced in the labor market, the observed causal effect of college refers to the college premium of at an older age. Therefore, the before- and after-expansion cohorts are not essentially comparable, because both the treatment and control groups refer to different population.
effect of college and the ability difference between high and low groups. Similarly, the college premium observed in post-expansion period for the always-takers as compared to never-takers (D_{2AN}) is a combination of the post-expansion treatment effect of college and the ability difference between high and low-ability groups. Taking the first difference in difference (DID_1), I have thus isolated the changing effect of college premium.

Furthermore, the difference between compliers and never-takers in pre-expansion period (D_{1CN}) can be conceptualized as the ability difference between median and low ability groups, whereas such difference (D_{2CN}) in post-expansion period is the combination of the post-expansion college treatment effect in addition to the effect produced by ability difference. Therefore, taking the second difference in difference (DID_2), I have isolated the post-expansion treatment effect of college. Knowing DID_1 and DID_2, naturally, we could arrive at the pre-expansion treatment effect of college.

Having obtained both the pre and post-expansion treatment effect of college, I then decompose the ability bias for the high vs. low ability group as well as for the median vs. low ability group. The first can be obtained via taking the difference between D_{1AN} (pre-expansion college premium for always-takers vs. never-takers) and the newly obtained pre-expansion treatment effect of college (DID_1 vs. DID_2), whereas the latter can be derived from taking the difference between D_{2CN} and DID_2. In doing so, the sources of the changing returns to college can be methodologically identified.

**RESULTS**

*Trends of Changing Returns to Education*
The first aim of this chapter is to depict the trends of returns to education in the reform periods. For this purpose, Table 6.6 summarizes the OLS estimation of educational returns and their 95% confidence intervals for years with available income information. Using a basic Mincer equation, the model only has a minimum set of covariates, so the estimated effects are only informative for the general trends of educational returns. Consistent with other studies (Fleisher and Wang 2005; Liu and Zhang 2008; Zhang et al 2005; Zhou 2000a), returns to education were increasing before 2002.

However, the trends have changed due to the 1999 massive educational expansion. Starting from 2003 when the first after-expansion college cohort began to enter the labor market, the yearly rate of schooling returns declines a bit during 2003-2008, from 8.8 percent to 8.6 percent. The decline in the college premium is more substantial: controlling for other variables, college graduates’ income on average is 66.7 percent higher than senior high school graduates in 2003, but the advantage declines to 49.8 percent in 2008. The figure rises to 73.8 percent in 2010, but according to the 2010 wave of the CHNS data, it shrinks again to 60.6 percent in 2011. In the meanwhile, junior college graduates’ income premium over senior high school graduates’ is generally stable after 2003. Given that the year 2003 marked the labor market entry year for people who entered college in 1999, it is evident that the influx of more skilled workers into the labor market coincided with the decreasing returns to college education at least in the first several years. Figure 6.4 depicts these trends.

The OLS estimation, however, does not take into accounts of the changing cohort quality and the changing composition of college graduates. As Rosenbaum (2003: 23) shows, the conventional estimates of the changing educational returns incorporate both the returns to unobserved ability and changes in the ability composition. To illustrate the shifts in the cohort
### Table 6.6: OLS Estimates of Returns to Education in Urban China, 1978–2010

<table>
<thead>
<tr>
<th>Year</th>
<th>Yearly Return (^a)</th>
<th>University vs HS (^b)</th>
<th>Junior College vs HS (^b)</th>
<th>Source</th>
<th>Retrospective? (^c)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1978</td>
<td>1.8 0.9 2.8</td>
<td>14.8 -0.6 30.2</td>
<td>12.6 (^\dagger) -2.4 27.6</td>
<td>State94</td>
<td>Yes</td>
</tr>
<tr>
<td>1984</td>
<td>1.5 0.5 2.5</td>
<td>18.9 4.0 33.9</td>
<td>-0.2 (^\dagger) -14.2 13.8</td>
<td>State94</td>
<td>Yes</td>
</tr>
<tr>
<td>1985</td>
<td>1.1 (^\ddagger) -0.1 2.2</td>
<td>14.0 -4.7 32.6</td>
<td>(-14.6) (^\dagger) -29.6 0.4</td>
<td>Change00</td>
<td>Yes</td>
</tr>
<tr>
<td>1987</td>
<td>2.8 1.9 3.7</td>
<td>20.1 9.0 31.2</td>
<td>5.3 (^\dagger) -6.5 17.1</td>
<td>State94</td>
<td>Yes</td>
</tr>
<tr>
<td>1988</td>
<td>2.2 2.0 2.5</td>
<td>17.9 14.6 21.2</td>
<td>7.4 4.8 10.1</td>
<td>CHIP88</td>
<td></td>
</tr>
<tr>
<td>1989</td>
<td>4.6 4.0 5.2</td>
<td>45.1 31.6 58.7</td>
<td>15.1 3.8 26.4</td>
<td>CHNS89</td>
<td></td>
</tr>
<tr>
<td>1990</td>
<td>2.0 1.6 2.3</td>
<td>17.1 13.0 21.2</td>
<td>9.2 5.9 12.5</td>
<td>CHIP95</td>
<td>Yes</td>
</tr>
<tr>
<td>1991</td>
<td>3.7 2.8 4.5</td>
<td>14.7 4.1 25.4</td>
<td>7.8 (^\dagger) -1.4 17.0</td>
<td>State94</td>
<td>Yes</td>
</tr>
<tr>
<td>1992</td>
<td>3.8 2.9 4.6</td>
<td>13.9 1.1 26.6</td>
<td>10.5 2.1 18.9</td>
<td>State94</td>
<td>Yes</td>
</tr>
<tr>
<td>1993</td>
<td>3.6 2.6 4.7</td>
<td>18.2 6.6 29.8</td>
<td>15.9 7.6 24.1</td>
<td>State94</td>
<td>Yes</td>
</tr>
<tr>
<td>1994</td>
<td>5.8 5.1 6.4</td>
<td>30.5 21.9 39.2</td>
<td>27.6 20.5 34.6</td>
<td>State94</td>
<td></td>
</tr>
<tr>
<td>1995</td>
<td>4.0 3.7 4.3</td>
<td>26.1 22.7 29.4</td>
<td>16.6 13.8 19.4</td>
<td>CHIP95</td>
<td></td>
</tr>
<tr>
<td>1996</td>
<td>4.4 3.5 5.2</td>
<td>36.6 24.8 48.4</td>
<td>15.2 4.7 25.7</td>
<td>Life96</td>
<td></td>
</tr>
<tr>
<td>1997</td>
<td>5.0 1.8 8.2</td>
<td>27.0 (-54.0) 108.0</td>
<td>34.8 19.3 50.3</td>
<td>CHNS97</td>
<td></td>
</tr>
<tr>
<td>1998</td>
<td>4.7 4.3 5.1</td>
<td>33.6 29.7 37.5</td>
<td>21.6 18.4 24.7</td>
<td>CHIP02</td>
<td>Yes</td>
</tr>
<tr>
<td>1999</td>
<td>5.2 4.8 5.6</td>
<td>39.2 35.4 43.1</td>
<td>24.9 21.7 28.1</td>
<td>CHIP02</td>
<td>Yes</td>
</tr>
<tr>
<td>2000</td>
<td>6.5 5.5 7.4</td>
<td>47.2 37.8 56.6</td>
<td>35.5 26.5 44.4</td>
<td>Change00</td>
<td></td>
</tr>
<tr>
<td>2001</td>
<td>6.1 5.7 6.5</td>
<td>45.7 41.7 49.6</td>
<td>27.8 24.7 30.9</td>
<td>CHIP02</td>
<td>Yes</td>
</tr>
<tr>
<td>2002</td>
<td>6.6 6.2 7.1</td>
<td>47.7 43.5 51.9</td>
<td>30.0 26.9 33.1</td>
<td>CHIP02</td>
<td></td>
</tr>
<tr>
<td>2003</td>
<td>8.8 7.7 10.0</td>
<td>66.7 56.1 77.3</td>
<td>40.7 32.4 48.9</td>
<td>CGSS03</td>
<td></td>
</tr>
<tr>
<td>2004</td>
<td>11.3 10.4 12.1</td>
<td>102.7 80.2 125.3</td>
<td>66.3 55.5 77.2</td>
<td>CHNS04</td>
<td></td>
</tr>
<tr>
<td>2005</td>
<td>8.3 7.5 9.1</td>
<td>58.2 47.5 68.9</td>
<td>43.7 35.2 52.2</td>
<td>CGSS05</td>
<td></td>
</tr>
<tr>
<td>2006</td>
<td>8.4 7.3 9.5</td>
<td>52.2 39.6 64.7</td>
<td>28.8 19.2 38.3</td>
<td>CGSS06</td>
<td></td>
</tr>
<tr>
<td>2008</td>
<td>8.6 7.2 10.0</td>
<td>49.8 32.7 66.9</td>
<td>42.2 27.9 56.6</td>
<td>CGSS08</td>
<td></td>
</tr>
<tr>
<td>2009</td>
<td>9.2 8.3 10.2</td>
<td>84.1 45.6 122.6</td>
<td>48.4 37.2 59.7</td>
<td>CHNS09</td>
<td></td>
</tr>
<tr>
<td>2010</td>
<td>9.8 8.9 10.7</td>
<td>73.8 63.1 84.5</td>
<td>41.9 32.7 51.1</td>
<td>CGSS10</td>
<td></td>
</tr>
<tr>
<td>2011</td>
<td>8.1 7.4 8.7</td>
<td>60.6 49.2 72.0</td>
<td>51.0 43.6 58.3</td>
<td>CHNS11</td>
<td></td>
</tr>
</tbody>
</table>

Notes: Reported are robust unstandardized coefficients and 95% confidence intervals (times 100). All the coefficients are significant at p<.05 level (two-tailed test) except for those marked with \(^!\). Data are weighted when weights are available (Life96 and CGSS03-10).

Except for the CHNS data, which does not have a variable for party membership, all the estimation is based on the model \(\text{Log}(Y) = a + b_1 \times \text{Sex} + b_2 \times \text{Age} + b_3 \times \text{Age-squared} + b_4 \times \text{Education} + b_5 \times \text{Party} + \text{Provincial dummies} + e\). Party membership is not included in the model for the CHNS data, which predicts larger returns. See also Liu and Zhang (2012) for their estimated educational returns based on the CHNS data.

\(^a\) “Education” is measured by years of schooling.

\(^b\) “Education” is measured by levels with “senior high school” as the reference group.

\(^c\) Retrospective data for yearly income or not.

Source: Table 6.6.
Note: Due to the lack of the party membership measure, the CHNS data predict higher educational returns than the prediction from other datasets. To avoid the bias, in this figure I do not plot the coefficients estimated from the CHNS data.

Figure 6.4: Rate of Yearly Schooling Return and the College Premium, Urban China, 1978-2010

quality and in the composition of college graduates, Figure 6.5 graphs the shares of educational groups by hukou origin.

As Figure 6.5 shows, the shares of educational groups differ significantly by urban and rural hukou origin. Among all those who have completed at least compulsory education, Figure 6.5a shows that the proportion of rural children who go to college is much smaller than their urban counterparts. This is because the limited opportunity at the senior high school level has prevented many rural children from obtaining college education. Figure 6.5b shows that among all the individuals who have at least completed senior high school, the share of college graduates has increased for the young cohorts who experience college transition during 1999-2010 (i.e. the
6.5a: Among worker who have completed junior high school education

Note: in both figures, the periods are captured by birth cohorts (Periods=birth cohorts+18).

Figure 6.5: The Changing Composition of Educational Groups, by Hukou Origin
after-expansion cohorts). The increase is greater for urban children than for rural children. In the second period, about 53.8 percent of urban children go to college; in the third period, the figure becomes about 76.2 percent, indicating that a substantive proportion of less-qualified urban children have been incorporated into college. Because of the increased heterogeneity in ability and in social origin background, the conventional OLS estimates fail to account for the heterogeneous treatment effects.

The Causal Effect of College on Earnings

I adopt the marginal treatment effect framework to understand the causal effect of college on earnings. To estimate MTE, I first run a logit model to obtain the propensity score of attending college. Table 6.7 presents the estimated logit model of college attendance, by periods of college transition. The variables included in the model are similar to those in Chapter 4, with the addition of mother’s years of schooling and ISEI scores as social origin variables, as well as a region and a birth cohort variable. To save space, the table only reports the linear effect of birth cohorts.

The results are similar to those presented in Chapter 4: social origin variables including hukou origin, father’s education and occupation are all significantly positive predictors of college attendance after the 1999 expansion. The interaction term between sex and hukou origin is also significant in the third period, suggesting a significant reversal of male’s advantage of college attendance in urban areas. Moreover, individuals from the eastern region on average have a larger probability to enter college than those from the central region. Therefore, geographic location is another significant determinant for the allocation of educational opportunities, in addition to the hukou origin variable.
Table 6.7: Estimated Logit Model for College Attendance, by Periods of College Transition, 1981-2010

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coefficients</td>
<td>Marginal effect</td>
<td>Coefficients</td>
<td>Marginal effect</td>
<td>Coefficients</td>
<td>Marginal effect</td>
</tr>
<tr>
<td><strong>Hukou origin (urban)</strong></td>
<td>0.779**</td>
<td>0.131**</td>
<td>0.596*</td>
<td>0.142*</td>
<td>1.134**</td>
<td>0.238**</td>
</tr>
<tr>
<td>(0.215)</td>
<td>(0.035)</td>
<td>(0.251)</td>
<td>(0.057)</td>
<td>(0.208)</td>
<td>(0.208)</td>
<td>(0.042)</td>
</tr>
<tr>
<td>Father’s schooling</td>
<td>0.018</td>
<td>0.036</td>
<td>0.099</td>
<td>0.064*</td>
<td>0.013*</td>
<td></td>
</tr>
<tr>
<td>(0.019)</td>
<td>(0.003)</td>
<td>(0.028)</td>
<td>(0.007)</td>
<td>(0.027)</td>
<td>(0.005)</td>
<td></td>
</tr>
<tr>
<td>Father’s schooling</td>
<td>0.061**</td>
<td>0.010**</td>
<td>-0.005</td>
<td>-0.001</td>
<td>-0.028</td>
<td>-0.006</td>
</tr>
<tr>
<td>(0.017)</td>
<td>(0.003)</td>
<td>(0.026)</td>
<td>(0.006)</td>
<td>(0.028)</td>
<td>(0.006)</td>
<td></td>
</tr>
<tr>
<td>Father’s ISEI</td>
<td>0.002</td>
<td>0.010*</td>
<td>0.002*</td>
<td>0.019**</td>
<td>0.004**</td>
<td></td>
</tr>
<tr>
<td>(0.005)</td>
<td>(0.001)</td>
<td>(0.005)</td>
<td>(0.001)</td>
<td>(0.007)</td>
<td>(0.001)</td>
<td></td>
</tr>
<tr>
<td>Mother’s ISEI</td>
<td>0.012*</td>
<td>0.021**</td>
<td>0.005**</td>
<td>0.007</td>
<td>0.001</td>
<td></td>
</tr>
<tr>
<td>(0.005)</td>
<td>(0.001)</td>
<td>(0.007)</td>
<td>(0.002)</td>
<td>(0.007)</td>
<td>(0.001)</td>
<td></td>
</tr>
<tr>
<td>Number of siblings</td>
<td>-0.027</td>
<td>-0.113</td>
<td>-0.027</td>
<td>0.110</td>
<td>0.022</td>
<td></td>
</tr>
<tr>
<td>(0.053)</td>
<td>(0.082)</td>
<td>(0.020)</td>
<td>(0.081)</td>
<td>(0.016)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sex (male)</td>
<td>0.481*</td>
<td>0.081*</td>
<td>0.655**</td>
<td>0.156**</td>
<td>0.367*</td>
<td>0.074*</td>
</tr>
<tr>
<td>(0.211)</td>
<td>(0.035)</td>
<td>(0.240)</td>
<td>(0.056)</td>
<td>(0.188)</td>
<td>(0.038)</td>
<td></td>
</tr>
<tr>
<td>Sex*Hukou origin</td>
<td>-0.256</td>
<td>-0.422</td>
<td>-0.100</td>
<td>-0.857**</td>
<td>-0.183**</td>
<td></td>
</tr>
<tr>
<td>(0.265)</td>
<td>(0.042)</td>
<td>(0.268)</td>
<td>(0.062)</td>
<td>(0.266)</td>
<td>(0.056)</td>
<td></td>
</tr>
<tr>
<td>Region^a</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>West</td>
<td>0.000</td>
<td>0.615**</td>
<td>0.151**</td>
<td>0.119</td>
<td>0.024</td>
<td></td>
</tr>
<tr>
<td>(0.174)</td>
<td>(0.220)</td>
<td>(0.054)</td>
<td>(0.251)</td>
<td>(0.049)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>East</td>
<td>0.624**</td>
<td>0.805**</td>
<td>0.192**</td>
<td>0.415*</td>
<td>0.084*</td>
<td></td>
</tr>
<tr>
<td>(0.164)</td>
<td>(0.208)</td>
<td>(0.048)</td>
<td>(0.198)</td>
<td>(0.040)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Year of birth</td>
<td>0.028</td>
<td>-0.049*</td>
<td>-0.012*</td>
<td>0.037+</td>
<td>0.007^</td>
<td></td>
</tr>
<tr>
<td>(0.022)</td>
<td>(0.023)</td>
<td>(0.006)</td>
<td>(0.019)</td>
<td>(0.004)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>-58.399</td>
<td>94.215*</td>
<td>0.142*</td>
<td>-74.717*</td>
<td>0.238**</td>
<td></td>
</tr>
<tr>
<td>(42.597)</td>
<td>(45.692)</td>
<td>(0.057)</td>
<td>(38.147)</td>
<td>(0.042)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Pseudo R-square 0.096 0.083 0.087
Log-likelihood -781.8 -533.8 -664.8
N 1,885 1,032 1,494

^Note: Robust standard errors in parentheses. Data are weighted. Within-province clustering effects are controlled.
^p<.10; * p<.05; ** p<.01 (two-tailed test)
^Refer to “Central region”.

6.6a: 1981-1992 College Transition Cohorts

6.6b: 1993-1998 College Transition Cohorts

6.6c: 1999-2010 College Transition Cohorts

Figure 6.6: The Distribution of the Propensity Score, by Periods of College Transition
Based on the estimated propensity score, I proceed to estimate the marginal treatment effect on the common support, for each period of college transition. Figure 6.6 depicts the common support of the propensity score for each college transition cohort. It shows that not only does the college and high school graduates differ significantly in the propensity score in each cohort, as indicated by the common support, but also the distribution of the common support changes across cohorts. After the expansion, both the treated and untreated have a higher tendency of college-going. Therefore, both the observed and unobserved heterogeneity as captured by \( p_l Z_i \) and \( U_{sl} \) in equation (3) are shifting after the 1999 policy.

I use a parametric-normal approach to estimate the marginal treatment effect. The variables entering the income equation include sex, experience, experience-squared, party membership, current ISEI score, current work unit, and region, so the variables remaining in the schooling equation are hukou origin, father’s and mother’s years of schooling and ISEI score, number of siblings and year of birth. I then graph the estimated marginal treatment effect over the unobserved heterogeneity parameter \( U_s \). Figure 6.7 summarizes the pattern. Figure 6.7a and 6.7b show that the treatment effect is negatively associated with the resistance to college attendance \( (U_s) \) for the first two periods, which means that the larger the propensity to go to college, the higher the marginal treatment effect is. This is a scenario of positive selection, as discussed in Carneiro and Lee (2007), and Carneiro et al (2011). In this scenario, individuals who are more likely to attend college are also the ones with higher treatment effect.

However, as illustrated in Figure 6.7c, the treatment effect has become positively associated with the resistance to college attendance for the cohort after the expansion. This demonstrates that relationship between the treatment effect and the propensity to attend college is a case of negative selection, as identified in Brand and Xie (2010). In other words, individuals who are
6.7a: 1981-1992 College Transition Cohorts

![Graph](image)

6.7b: 1993-1998 College Transition Cohorts

![Graph](image)

6.7c: 1999-2010 College Transition Cohorts

![Graph](image)

Figure 6.7: MTE as a Function of Unobserved Heterogeneity ($U_s$): Parametric-normal Approach
less likely to attend college would benefit more, if they attended. The means that the negative selection bias may indicate involuntary selection: the unschooled groups would have benefitted more from schooling, yet there exist structural constraints that prevent them from attending college. Based on the findings from Chapter 4, the “bottleneck structure” has been constricted for socially disadvantaged groups, which would partially explain the negative selection pattern found here. To put differently, it is educational expansion that has changed the selection processes.

Table 6.8 further presents the estimated various treatment effect parameters, using the weights provided by Heckman and Vytiacil (2001, 2006a, 2006b). The instrument for the IV approach is the estimated propensity score. According to the estimated effect parameters, the income gap between college and high school graduates has become smaller for the after-expansion transition cohorts. Comparing the ATE and ATT parameters in the second and third periods, we further find that while ATT is higher for urban children than for rural children in the second period, the pattern is reversed in the third period. Moreover, there are more declines in the average treatment effect for urban children than for rural children after the expansion. One interpretation is that the college premium for urban children is dragged down by the influx of less-qualified urban high school graduates into college after the expansion. In other words, the average ability declines more for the urban college cohorts than for the rural college cohorts after the expansion. Again, this is because the existing “bottleneck” structure only selects the high-ability rural children into college, even after the college expansion (see Figure 6.5a).

The decomposition of the total bias further shows that both urban and rural children’s unobserved abilities are less rewarded in the third period. The sizes of the selection bias are both close to 0 after the expansion, suggesting that college graduates on average would not earn more
Table 6.8: Comparisons of Different Treatment Effect Parameters, by Periods of College Transition and Hukou Origin, 1981-2010

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total</td>
<td>Urban</td>
<td>Rural</td>
<td>Total</td>
<td>Urban</td>
<td>Rural</td>
</tr>
<tr>
<td>N</td>
<td>1681</td>
<td>957</td>
<td>352</td>
<td>936</td>
<td>584</td>
<td>352</td>
</tr>
<tr>
<td>OLS</td>
<td>.564</td>
<td>.418</td>
<td>.677</td>
<td>.644</td>
<td>.553</td>
<td>.678</td>
</tr>
<tr>
<td></td>
<td>(.048)</td>
<td>(.054)</td>
<td>(.105)</td>
<td>(.058)</td>
<td>(.068)</td>
<td>(.105)</td>
</tr>
<tr>
<td>IV</td>
<td>.623</td>
<td>.472</td>
<td>.680</td>
<td>.656</td>
<td>.588</td>
<td>.568</td>
</tr>
<tr>
<td></td>
<td>(.080)</td>
<td>(.106)</td>
<td>(.097)</td>
<td>(.094)</td>
<td>(.116)</td>
<td>(.162)</td>
</tr>
<tr>
<td></td>
<td>(.064)</td>
<td>(.075)</td>
<td>(.091)</td>
<td>(.059)</td>
<td>(.092)</td>
<td>(.117)</td>
</tr>
<tr>
<td>ATT</td>
<td>.309</td>
<td>.291</td>
<td>.427</td>
<td>.481</td>
<td>.487</td>
<td>.427</td>
</tr>
<tr>
<td></td>
<td>(.073)</td>
<td>(.087)</td>
<td>(.152)</td>
<td>(.075)</td>
<td>(.089)</td>
<td>(.157)</td>
</tr>
<tr>
<td>ATUT</td>
<td>.387</td>
<td>.201</td>
<td>.536</td>
<td>.465</td>
<td>.565</td>
<td>.536</td>
</tr>
<tr>
<td></td>
<td>(.061)</td>
<td>(.085)</td>
<td>(.122)</td>
<td>(.071)</td>
<td>(.126)</td>
<td>(.118)</td>
</tr>
</tbody>
</table>

Total Bias = OLS–ATE
Selection bias = OLS–ATT
Sorting gain = ATT–ATE

Note: The IV regression uses the propensity score as the instrument. The treatment effect estimates are based on the propensity score matching results. Standard errors shown in parentheses are obtained by bootstrapping. All the effects are statistically significant at the p<.05 level.

than senior high school graduates if they had not attended college. The selection effect is reduced more for the rural children. In the first two periods, the size of the selection bias is rather salient for rural children. The positive effects suggest that if the rural college graduates have not attended college, they would on average earn about 25 percent more than an average rural senior high graduate earns. The disappearance of the selection bias in the third period therefore suggests that the unobserved ability is negated in the labor market, particularly for rural children. As it doesn’t matter whether a high school graduate’s ability is high or low for earnings after the
expansion, not having a college credential becomes a bad signal in the labor market. The implication, to put crudely, is that if everyone goes to college, then one need to go to.

Finally, the null sorting effect found in the third period indicates that a college graduate’s earnings would not change significantly if he had switched to the high school track. This implies that the expansion has absorbed a certain group of individuals who would not particularly benefit from going to college.

The Policy Effect of Educational Expansion on the College Premium

The previous section has identified the causal effect of college on earnings in a given period. I now turn my attention to the policy effect of educational expansion on the changes in the college premium. Considering that the policy effect does not apply to everyone, I take a difference-in-difference approach to understand how the causal effect of college education differs among always-takers and compliers across cohorts. The method used to identify the compliers is described in the methodology section.

Table 6.9 presents the OLS estimates of the returns to college education for the 1993-1999 and 1999-2010 college transition cohorts. Considering the differences in the composition among senior high school and college graduates, the analyses in this section are limited to those in the common-support of the propensity scores (as shown in Figure 6.6b and 6.6c). To generate comparable results, again only a minimum set of independent variables is used in the models. They include sex, experience, party membership, a policy dummy variable (1993-1998 cohorts=0), and a group variable with 1=never –takers, 2=compliers, and 3=always-takers. Note that never-takers are senior high school graduates and always-takers are college graduates. The
Table 6.9: OLS Estimates of the Returns to College under Educational Expansion, within the Common Support, 1993-2010

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sex 0.299** 0.193 +</td>
<td>Sex 0.203** 0.176** 0.247*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.076) (0.096)</td>
<td>(0.053) (0.039) (0.102)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Experience 0.054 + 0.111**</td>
<td>Experience 0.052** 0.106** 0.030 +</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.032) (0.019)</td>
<td>(0.017) (0.016) (0.017)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Experience-square/100 -0.212 -0.196**</td>
<td>Experience-square/100 -0.126 -0.449** -0.007</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.122) (0.064)</td>
<td>(0.099) (0.090) (0.046)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Party membership 0.070 0.298**</td>
<td>Party membership 0.138* 0.130 0.167</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.099) (0.082)</td>
<td>(0.053) (0.075) (0.083)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Groups a Compliers 0.356* 0.413**</td>
<td>Compliers 0.335** 0.233 0.207</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.130) (0.106)</td>
<td>(0.104) (0.123) (0.208)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Always-takes 0.813** 0.496**</td>
<td>Always-takes 0.833** 0.775** 0.715**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.120) (0.143)</td>
<td>(0.123) (0.171) (0.139)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Policy 0.402** 0.358* 0.374*</td>
<td>Policy 0.402** 0.358* 0.374*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.115) (0.154) (0.140)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interaction</td>
<td>Compliers*Policy -0.001 0.153 0.065</td>
<td>Compliers*Policy -0.001 0.153 0.065</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.155) (0.179) (0.265)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Always-takes<em>Policy -0.436** -0.371</em> -0.569**</td>
<td>Always-takes<em>Policy -0.436** -0.371</em> -0.569**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.087) (0.176) (0.206)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Constant 6.429** 6.515**</td>
<td>Constant 6.486** 6.460** 6.498**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.221) (0.106)</td>
<td>(0.142) (0.129) (0.207)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>N 924 756</td>
<td>N 1,680 1,049 631</td>
<td></td>
</tr>
<tr>
<td></td>
<td>R-squared 0.159 0.131</td>
<td>R-squared 0.128 0.124 0.111</td>
<td></td>
</tr>
</tbody>
</table>

Note: Robust standard errors in parentheses. Data are weighted. Within-province clustering effects are controlled. All the estimates are within the common support of the matched propensity as shown in Figure 6.6b and 6.6c. + p<.10; * p<.05; ** p<.01 (two-tailed test)

a Refer to “Never-takers”.

Compliers are senior high school graduates before the expansion but would be college graduates after the expansion.

The first panel of Table 6.9 runs the estimation separately for the pre- and post-expansion college transition cohorts. It shows that the always-takers and compliers have significantly higher returns than never-takers for both cohorts. However, as there are composition differentials in ability between the two cohorts, I decompose the “pure” college premium for both cohorts and
returns to differential abilities based on the framework in Table 6.5. Since the coefficients are all significant for the always-takers and compliers in the two cohorts, I write that:

\[
\begin{align*}
\text{College}_2 - \text{College}_1 &= 0.496 - 0.831; \\
\text{College}_2 &= 0.413 - 0.356 \\
\text{College}_2 + \text{Return to high ability} &= 0.496 \quad (9) \\
\text{College}_2 + \text{Return to median ability} &= 0.413
\end{align*}
\]

where College_1 and College_2 refer to the “pure” premium associated with the college credential. Solving (9), I get that College_1 =0.392, College_2=0.057, Ability bias for the always-takers=0.439, and Ability bias for the compliers=0.356. Therefore, the “pure” college premium is 39.2 percent for the pre-expansion cohort and only 5.7 percent for the post-expansion cohort.

Returns to unobserved ability, however, is the main component to the observed income gaps between college and high school graduates after the expansion: it accounts for 88.4 percent of the observed income gap between always takers and never-takers and 86.2 percent of the observed income gap between compliers and never-takers, respectively. As a result, the changes in the college premium after the expansion are mainly driven by changes in the ability composition of the college graduates.

Next, I study how the policy effect differs by hukou origin. Results in the second panel shows that the college premium is declining for the always-takers, and the decline is more salient among rural children than their urban counterparts. Net of others, rural always-takers’ advantage over rural high school graduates has declined by 56.9 percent, whereas as urban always-takers’ advantage over urban high school graduates only declines by 37.1 percent. As a result, after the expansion, rural children of top percentiles only enjoy a 14.6 (71.5-56.9) percent of college premium over their rural high school classmates; in contrast, top urban children’s college premium is 40.0 (77.5-37.1) percent compared to their urban high school peers. Clearly, the 1999
college expansion has reduced the previous advantage enjoyed by the highly-qualified rural children (always-takers).

Moreover, the insignificant coefficient of the interaction term Compliers*Policy suggests that the income gap between compliers and never-takers does not significantly differ before and after the expansion. Thus, although the compliers become college attendees after the expansion, they do not significantly improve their relative earnings. In this regard, college education does not help the complier to get ahead in earnings.

SUMMARY AND CONCLUSION

Existing studies in the tradition of labor economics have focused on the return to education across time. Much research has focused on decomposing the effects of the changing college premium. Previous scholars have conceptualized three different sources of the effects of the changing college premium: i.e. the price effect, the signaling effect of unobserved ability and the composition effect due to the changed ability distribution for pre and post expansion cohorts. However, the exact source of the effect is subject to much debate.

The lack of consensus aside, the key limitation in the existing literature on the declining college premium is that sufficient attention has yet to be paid to the social origin heterogeneity with respect to the changing college premium. In other words, given that the college premium is changing after educational expansion, how does individuals’ social origin factors intervene such process? Do individuals with more disadvantageous social positions also stand to lose more?

This question is of great sociological importance. As the acquiring of income is an indispensable step in the association of social origin, education and destination, understanding the social origin heterogeneity and answering the question of who have benefited more or
suffered less when college premium is declining allow us to fully understand the implications of educational expansion on the lasting patterns of education-based social stratification and inequality.

Therefore, in this chapter, I turn my focus to the social origin heterogeneity in the declining college premium after educational expansion. Based on the existing literature, particularly building on the research that aims to uncover the sources of effects of declining returns to education, I formulated two specific hypotheses.

Using a variety of national representative survey data on reforming China, I come to the following four main findings. First, college premium and yearly return to education has been declining since 2003, the year in which post-expansion college graduates have begun to enter the labor market and cause sharp influx of larger quantity of workers with higher education.

Second, I find that there is a reversed selection pattern across reform periods. Whereas before educational expansion, there is a positive selection in the economic returns to college (Carneiro et al 2010, 2011), the pattern becomes negative selection after the expansion (Brand and Xie 2010; Tsai and Xie 2008). The change in the selection patterns suggests that educational expansion has strengthened the structural constraint for individuals with high propensity of college attendance but who could not go to college. Due to China’s “bottleneck” structure, these people could be individuals with high ability from rural areas.

Third, I find that the college premium declines more after the expansion for rural children in the top percentiles than for their urban counterpart. The advantage of these highly-qualified rural children (always takers) over the rural high school graduates shrinks by 57 percent after the expansion. The returns to their unobserved ability have also declined after the expansion. It is thus the rural children of the top percentiles that fall behind in the college premium under
educational expansion. In addition, I find that the compliers do not really benefit from educational expansion in improving their relative earnings.

Forth, my decomposition of the changing college premium shows that the “true” college premium, or the real price of college education, is rather small for the after-expansion cohort. Instead, much of the change in the college premium is due to changes in the composition of ability among the college graduates. Returns to ability are also heterogeneous among the college graduates with high- and median-levels of ability.

These findings provide solid support for my two hypotheses. Taken together with the findings in Chapter 4 and 5, these results demonstrate that under China’s 1999 educational expansion, urban children not only “benefit first and most” in terms of educational and occupational attainment, they also “suffer less” when it comes to the declining college premium. By contrast, the top percentiles of the rural children fall behind the most in terms of college premium. Further analyses are needed to test whether individuals of lower class could still benefit from going to college because of the negative selection pattern under educational expansion.

In a word, based on my theoretical framework in Chapter 2, I argue that the observed increasing education-based income inequality in the labor market is a result of educational expansion itself. With the sudden increase of educational opportunities at the tertiary level, the advantageous group’s previous advantages in the educational attainment process can be transmitted into the labor market earnings distribution (Willis and Rosen, 1979; Heckman et al 1996). In this regard, the transmission of social advantages/disadvantages is a cumulative and asymmetric process (Xie and Wu 2005). Coupled with the shifting population composition
brought forth by educational expansion, fully understanding the changes in inequality patterns in labor market outcomes cannot separate out the dynamic processes of educational expansion.
Chapter 7

Conclusions

Conventional wisdom suggests that with long-term growth of school enrollments, education will become an increasingly important mechanism for the transmission of social status. That is, along the process of technological advance and economic development, there tends to be greater equality of opportunity with respect to both educational and occupational attainment. In this dissertation, I have examined how educational expansion affects education-based inequality and mobility, using reforming China as an illustrative case. The general empirical puzzle that I address is: why there exist various patterns of origin-education-destination associations among different societies under educational expansion.

After the introduction provided in Chapter 1, in Chapter 2, I developed formal modeling towards a generalized theory of education-based inequality and mobility. I theorize that educational-based inequality and mobility is affected by educational expansion sequentially. That is, the inequality pattern for one post-compulsory level of education is linked to that of the subsequent level. In a given society, the effect of educational expansion on education-based inequality and mobility is contingent upon three factors: (1) the initial social group differences in the ability and resource distribution, (2) the structural and (3) the institutional changes introduced by educational expansion that modify the ability and resource requirements for schooling. Across societies, institutional differences in the sociopolitical contexts further drive the variations in the patterns caused by the three factors. Furthermore, I theorize that educational expansion can have a compositional effect on the origin-education-destination associations and on income returns to education.
To test this generalized theory, I turn my attention to reforming China as an illustrative case. In Chapter 3, using a variety of national statistics, I introduced the historical, economic and social context of China’s reforms and educational expansion, while highlighting three aspects of consequences of educational expansion. China’s massive educational expansion since 1999 is characterized by an increasingly constrained “bottleneck” structure and rising educational costs for senior high school and college education. These structural and institutional changes under educational expansion have affected not only how educational opportunity is allocated, but also how education is rewarded in the labor market. Thus, to test the general theory of Chapter 2, using the specific Chinese case, I have hypothesized that educational expansion would enlarge educational inequality at the college level, decrease the labor market returns to college education, and strengthen the effect of social origin on the labor market returns to college.

Based on these hypotheses, Chapter 4, 5 and 6 each tackles one aspect of the consequences of educational expansion. In Chapter 4 I ask, who have gotten ahead in educational attainment under the expansion? Using the pooled data from five highly comparable national representative samples, I investigate changes in the social origin effects on educational transitions so as to understand the changing patterns of educational stratification during China’s economic reform era and under educational expansion from 1981 to 2010. My findings have shown that although urban-rural differentials and social origin differentials in access to vocational school both decreased in the 1990s, such differentials in the transition to senior high school increased because of contracted high school opportunity in this period. After the 1999 expansion, however, while educational inequality in going to senior high school was generally declining, the allocation of college opportunity became less equal. Urban-rural inequality in the transition to college increased, and father’s social origin status began to play a more important
role in determining one’s likelihood of entering college. Therefore, to put it simply, the rapid expansion of higher education in China since 1999 has mainly benefited urban children and children from better-off families (e.g., sons and daughters of managers/professionals).

Moving beyond school, the relationship between education and first occupation is an important issue in sociology. In Chapter 5 I ask, who have risen higher in terms of first occupation attainment under China’s educational expansion? By using multiple measurements of labor market entry positions, my empirical analyses suggest that China’s stratified educational system functions as a sorting and allocating mechanism. Particularly, higher educational credentials confer a higher return in occupational status. In addition, secondary vocational education also gives individuals a higher return of status in first jobs than senior high school education. However, as compared to the 1980s, father’s occupation has become a more important determinant for children’s occupational attainment in the 1990s, controlling for education. Similar to the findings of Chapter 4, here, in the context of first occupation attainment, my findings suggest that it is individuals from advantaged social origin positions who have benefited more under the educational expansion. In other words, education as a systematic mobility channel has lost its significance for individuals from disadvantaged social origin backgrounds.

Having examined the social origin differentials in educational and first occupation attainment under China’s educational expansion, I shift my attention to yet another indispensable step in the association of social origin, education and destination: the acquiring of income. In Chapter 6 I ask, who has fallen behind in the college premium under educational expansion? I turn my focus to the social origin heterogeneity in the declining college premium after educational expansion. Using a variety of nationally representative survey data on reforming China, I have found that first, since 2003, i.e. the year when post-expansion graduates began to
enter into the labor market, the college premium and yearly return to education have started
decreasing. My decomposition of the changing college premium shows that the “true” college
premium, or the real price of college education, is rather small for the after-expansion cohort.
Instead, much of the change in the college premium is due to changes in the composition of
ability among the college graduates. In addition, there is a reversed selection pattern (i.e. from
positive to negative selection) during the reform periods. Such change suggests that educational
expansion has strengthened the structural constraint for individuals with high propensity of going
to college yet could not attend. Due to China’s “bottleneck” structure and drawing on the
findings from Chapter 4, I argue that these are the individuals with high ability from less-
advantaged social origin backgrounds (e.g. individuals from rural areas). Furthermore, in the face
of the general decline of college premium after the expansion, such declines are more
pronounced for rural children in the top percentiles than for their urban counterpart. The
advantage of these highly qualified rural children over the rural high school graduates shrunk by
57 percent after the expansion. The returns to their unobserved ability have also declined after
the expansion. It is thus the rural children of the top percentiles that fall behind in the college
premium under educational expansion.

Taken together with the findings in Chapters 4 and 5, and tying back to the general theory
of education-based inequality proposed in Chapter 2, overall my dissertation has demonstrated
that under China’s 1999 educational expansion, urban children, and individuals of advantaged
social origins not only have “benefit first and most” in terms of educational and occupational
attainment, they have also “suffered less” when it comes to the declining college premium. By
contrast, the top percentiles of the rural students, and individuals of disadvantaged social origins,
have fallen behind the most in terms of educational attainment, first occupation attainment as well as college premium.

Although abundant scholarly work on how educational expansion affects the patterns of education-based inequality and mobility exists, the field has remained theoretically and empirically inconclusive. Exiting studies have largely ignored the fact that educational expansion produces multifaceted consequences that are intrinsically connected. Educational expansion affects not only how different social groups make use of educational opportunities, but also, as students graduate and enter the labor market, the expansion in turn implicate how various social groups receive occupational and material returns to educational credentials. It is therefore paramount to integrate the various consequences of educational expansion. Otherwise, it is difficult to fully understand how educational expansion shapes the allocation of educational and occupational opportunities as well as the distribution of material resources in the labor market.

In addition, comparative research has found both similarities and variations in the patterns of education-based inequality and mobility (Breen and Luijx 2004a; Erikson and Goldthorpe 1992). Although there is an increasing consensus in the literature that industrialized (and post-industrial) countries share different education-based stratification patterns as compared to countries like Russia and China (Breen 2004; Breen and Jonsson 2005; Geber and Hout 1995, 1998, 2004; Guo and Wu 2009; Wu 2010; Wu and Treiman 2007), little is known about whether and to what extent this “empirical puzzle” holds in the context of educational expansion. Furthermore, on the theoretical front, the explanations of the impacts of educational expansion have long been fragmented and under-theorized. We lack a coherent theory that incorporates the various accounts, both micro and macro-based, in the existing literature. The micro-level rational choice approach focuses on how students with different social backgrounds make decisions for
educational attainment and social mobility whereas the macro-level institutional and structural approaches tend to emphasize the effects on social stratification and inequality patterns of the institutional and structural changes. Yet, despite the importance of the micro-macro linkage in sociology, the dialogue between the two approaches has remained extremely limited.

Finally, and perhaps most importantly, partially due to the missing micro-macro linkage in the literature, current research has yet to fully explain the variations found in the patterns of education-based inequality and mobility across societies. To be specific, none of the existing theories formulated based on the experience of industrialized societies could convincingly incorporate the “outlier patterns” observed in certain industrializing societies such as China and Russia. Given that the institutional and structural changes introduced by educational expansion differ by societal settings, and the impacts of educational expansion on education-based inequality and mobility patterns also differ accordingly, a thorough demonstration and careful explanations of the experience of industrializing countries such as China is called for.

Therefore, in this dissertation, I make both theoretical and empirical contributions to the existing scholarship of inequality, stratification and mobility. On the theoretical front, I have formalized a general theory of education-based inequality that accounts for both micro and macro factors. On the empirical front, I have examined the consequences of educational expansion from a dynamic perspective, focusing on the three aspects of consequences, i.e. educational attainment, school-to-work transition, and college premium. Contributing the general comparative research, I have provided a solid description of the trends, patterns and mechanisms of how educational expansion affects education-based inequality and mobility in the case of reforming China. Now, looking back at the question I raised at the beginning of this dissertation, I have found that unlike what might be suggested by the conventional wisdom, individuals from
advantaged social backgrounds have benefitted more every step of the way. These continued and cumulative advantages are shaped by not only the initial social group differences in the ability and resource distribution, but also the structural and the institutional changes introduced by educational expansion, which differ across societies as the structural and institutional arrangements in the sociopolitical contexts diverge. Therefore, with this dissertation, I call for a greater emphasis on understanding and theorizing the cumulative and asymmetrical social process that shapes the patterns and mechanism of education-based social stratification and inequality.
## Appendix

Table A1: Measurement of Main Variables in the Five Surveys

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Educational transitions</td>
<td>Educational history</td>
<td>No history</td>
<td>No history</td>
<td>Educational history</td>
<td>No history</td>
</tr>
<tr>
<td>Year of transition</td>
<td>Educational history</td>
<td>End year</td>
<td>End year</td>
<td>Educational history</td>
<td>End year</td>
</tr>
<tr>
<td>Hukou</td>
<td>Current</td>
<td>Current</td>
<td>Current</td>
<td>Current</td>
<td>Current</td>
</tr>
<tr>
<td>Hukou origin</td>
<td>When 14</td>
<td>No history</td>
<td>When 14 &amp; current, all biased</td>
<td>Year of change</td>
<td>Year of change</td>
</tr>
<tr>
<td>Father’s hukou</td>
<td>When 14, no current</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sibling</td>
<td>Not biased</td>
<td>Biased</td>
<td>Not biased</td>
<td>Not biased</td>
<td>Biased</td>
</tr>
<tr>
<td>Father’s education</td>
<td>Current</td>
<td>When 14</td>
<td>Current</td>
<td>Current</td>
<td>Current</td>
</tr>
<tr>
<td>Mother’s education</td>
<td>Current</td>
<td>When 14</td>
<td>Current</td>
<td>Current</td>
<td>Current</td>
</tr>
<tr>
<td>Father’s EGP</td>
<td>When 14</td>
<td>When 14</td>
<td>When 18</td>
<td>When 14</td>
<td>When 14</td>
</tr>
<tr>
<td>Father’s ISEI</td>
<td>When 14</td>
<td>When 14</td>
<td>When 18</td>
<td>When 14</td>
<td>When 14</td>
</tr>
<tr>
<td>Father’s work unit</td>
<td>When 14</td>
<td>When 14</td>
<td>When 18</td>
<td>When 14</td>
<td>When 14</td>
</tr>
<tr>
<td>Father’s party</td>
<td>Current and When 14</td>
<td>Current</td>
<td>Current and When 14</td>
<td>Current</td>
<td>Current</td>
</tr>
<tr>
<td>First job year</td>
<td>History</td>
<td>Yes, no history</td>
<td>Yes, no history</td>
<td>History</td>
<td>No information</td>
</tr>
<tr>
<td>First job work unit</td>
<td>History</td>
<td>No information</td>
<td>Yes, no history</td>
<td>History</td>
<td>No information</td>
</tr>
<tr>
<td>First job education</td>
<td>Yes</td>
<td>Yes, biased</td>
<td>Yes, biased</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>First job party</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>First job ISEI</td>
<td>History</td>
<td>No history</td>
<td>History</td>
<td>History</td>
<td>No information</td>
</tr>
<tr>
<td>Current ISEI</td>
<td>History</td>
<td>No history</td>
<td>History</td>
<td>History</td>
<td>Yes</td>
</tr>
<tr>
<td>Current EGP</td>
<td>History</td>
<td>No history</td>
<td>History</td>
<td>History</td>
<td>Yes</td>
</tr>
<tr>
<td>Income</td>
<td>Yearly</td>
<td>Yearly</td>
<td>Yearly</td>
<td>Yearly</td>
<td>Yearly</td>
</tr>
</tbody>
</table>
Table A2: List of Main Datasets Used (CHNS Data Not Included)

<table>
<thead>
<tr>
<th>Datasets</th>
<th>Year individual income available</th>
<th>Population age</th>
</tr>
</thead>
<tbody>
<tr>
<td>State and Life Chances 1994</td>
<td>1978, 84, 87, 91, 92, 93, 94</td>
<td>20-65, Urban</td>
</tr>
<tr>
<td>Life Histories 1996</td>
<td>1996</td>
<td>18-69</td>
</tr>
<tr>
<td>CHIP 1988</td>
<td>1988</td>
<td>Resident population</td>
</tr>
<tr>
<td>CHIP 1995</td>
<td>1995, 94, 93, 92, 91, 90</td>
<td>Resident population</td>
</tr>
<tr>
<td>CGSS 2003</td>
<td>2002</td>
<td>18-69</td>
</tr>
<tr>
<td>CGSS 2005</td>
<td>2005</td>
<td>18-69</td>
</tr>
<tr>
<td>CGSS 2006</td>
<td>2006</td>
<td>18-69</td>
</tr>
<tr>
<td>CGSS 2008</td>
<td>2010</td>
<td>18-69</td>
</tr>
<tr>
<td>CGSS 2010</td>
<td>2010</td>
<td>18-69</td>
</tr>
</tbody>
</table>
Bibliography


Ganzeboom, Harry BG, Ruud Luijkkx, and Donald J Treiman. 1989. "Intergenerational class mobility in comparative perspective " *Research in Social Stratification and Mobility* 8:3-84.


238


Hu, Anning, and Jacob Hibel. 2013. "Increasing heterogeneity in the economic returns to higher education in urban China." *The Social Science Journal*.


