Three Essays on the Higher Education Expansion in China

Qiao Wen

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
under the Executive Committee
of the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2020

Abstract

Three Essays on the Higher Education Expansion in China

Qiao Wen

My dissertation intends to better understand the impacts of large-scale education expansion programs on students' education and labor market outcomes both by reviewing related theory and prior literature, and by empirically analyzing a radical and large-scale higher education expansion program initiated in 1999 in China.

In Chapter 1, I review theories, methods and empirical studies on the labor market consequences of education expansion from both the partial equilibrium treatment effect and general equilibrium structural model literature. This chapter serves as the theoretical and methodological foundation for my later analyses in Chapter 2 and Chapter 3, and provides motivation for my empirical work because prior literature has not reached a consensus in terms of the impacts of large-scale education expansion programs on individuals' labor market outcomes or the wage structure in the labor market.

In Chapter 2, I take advantage of the fact that the substantially expanded access to higher education after China's higher education expansion provides plausibly exogenous variation in the probability of college attendance for students of different cohorts and coming from different

provinces. I thus employ a two-way fixed-effect model to estimate the expansion's causal impacts on individuals' education and labor market outcomes, and find that the expansion substantially improved educational outcomes, such as years of schooling completed, the probabilities of attending college and obtaining any post-secondary degree. The expansion also increased treated individuals' probability of working and earning positive income, and modestly improved their hourly income. However, the expansion's earnings effects are less robust to the exclusion of two largest metropolitan cities in China and the inclusion of province-year-level time-varying covariates to control for potential cofounding influences.

In Chapter 3, I exploit multiple repeated cross-sections of data to explore how the expansion affects the labor market at large, especially the college-high school earnings gap. Incorporating an aggregate labor supply model with imperfect substitution across labor with the same education level but in different age groups, I decompose the changes in age-group specific college premium over time into changes in the aggregate and cohort-specific relative supply of college-educated (vs. high school-educated) labor, and in the aggregate relative demand for college-educated labor. My findings show that a 1 percent increase in the relative supply of BA-educated workers within one's own cohorts would depress the BA-HS wage gap by 0.04 percentage point. Given that college enrollment increased by nearly 4 times from 1998 to 2005, the negative cohort effects could be substantial: for example, the cohort-specific relative supply for the youngest age group in my analysis increased by 112 percent from 2002 to 2009, suggesting an additional 4.5 percentage points decrease in the BA-HS wage gap for workers of this particular age group, on top of the effects of changes in aggregate relative supply and demand that are borne by workers in all age groups. Moreover, my estimates reveal a steadily increasing relative demand for BA-educated labor that raises college premium by approximately

2-3 percentage points annually; it is mitigated by the negative effects from the increase in the aggregate relative supply of BA-educated labor though; the latter effect also implies that the expansion has negative spillover effects on workers who attended college before the expansion.

Putting together, my dissertation provides a holistic picture of the full impacts of one of the largest education expansion program on record. My work is among the first to systematically analyze how the expansion affects "treated" individuals and the labor market at large, and therefore could contribute to all levels of decision-making. Findings from my analyses could also have global implications for much broader issues such as education-related income inequality, and the general equilibrium and distributional effects of large-scale social programs.

Table of Contents

List of Figures iv

List of Tables v

Acknowledgements vi

Dedication vii

Chapter 1: Estimating Labor Market Effects of Large-scale Education Programs: A Review of the Literature with Application to China's Higher Education Expansion 1

1.1 Introduction 1

1.2 Using Large-scale Education Programs as Quasi-experiments 4

1.3 Structural Models of Labor Market Dynamics 10

1.3.1 Shifts of the aggregate and cohort-specific supply curves 13

1.3.2 Shifts of the demand curve 14

1.3.3 Changes in quality composition of different types of labor 17

1.4 Recent Studies Bridging the Two Approaches 19

1.5 The Chinese Context 22

1.5.1 Returns to education in modern China 23

1.5.2 Higher education and the higher education expansion in China 24

1.5.3 Empirical studies on the causal effects of the higher education expansion on labor market outcomes 31

1.6 Summary 35

Chapter 2: Estimating Causal Returns to College: Quasi-experimental Evidence from China's Higher Education Expansion 37

2.1 Introduction 37
2.2 The 1999 Higher Education Expansion in China 40
2.3 Data and Method 44
2.3.1 Data 44
2.3.2 Sample 45
2.3.3 Estimation Strategy 47
2.3.4 Key Variables 48
2.4 Results 57
2.4.1 The College Expansion's Effects on Education 59
2.4.2 The College Expansion's Effects on Labor Market Outcomes 61
2.4.3 Estimating the Causal Returns to College 63
2.5 Falsification Tests and Robustness Checks 67
2.5.1 Falsification Tests 67
2.5.2 Robustness Checks 72
2.5.3 Alternative Channels for the Expansion's Effects 78
2.6 Discussion 80
2.6.1 Limitation of Using One Cross-section of Data 80
2.6.2 Limitation of Estimating Treatment Effects When General-Equilibrium Effects Present 81
2.7 Summary 83
Chapter 3: Evolution of the College Earnings Premium for Younger and Older Workers during the 2000s in China 85

3.1 Introduction 85
3.2 Evolution of College Earnings Premium during 2002-2009 86
3.3 Theoretical Framework 91
3.4 Data and Method 93
3.4.1 Data 94
3.4.2 Sample 95
3.4.3 Key Variables 96
3.4.4 Estimating the Elasticity of Substitution between Age Groups and Efficiency Parameters for Different Age Groups 102
3.5 Decomposing the BA-HS Wage Gap 104
3.6 Discussion and Conclusion 110
References 114
Appendix A: Appendix Figures and Tables 126
Appendix B: Additional Notes on Robustness Checks for Chapter 2 140

List of Figures

Figure 2.1: Undergraduate Enrollment in and Graduates from Regular Degree-granting Higher Education Institutions in China (in millions), 1995-2018. 42

Figure 2.2: College Admission Ratio for Selected Provinces, Including and Excluding Beijing and Shanghai, 1995-2008. 52

Figure 3.1: BA-HS Earnings Gap by Year for the Worker Sample and by Selected Age Groups. 87

Figure 3.2: Number of Graduates from 3- and 4-year Regular Higher Education Institutions by Year. 88

Figure 3.3: Age-group Specific Relative Supply Index for BA- vs. HS-educated Labor by Year. 99

List of Tables

Table 2.1: Descriptive Statistics for the Full and Worker Samples. 57

Table 2.2: Estimating the Effects of the Expansion on Education Outcomes. 59

Table 2.3: Estimating the Effects of the Expansion on Labor Market Outcomes. 62

Table 2.4: 2SLS Estimates of Expansion-induced Increases in Education on Labor Market Outcomes. 64

Table 2.5: Falsification Tests - DiD Estimators of the Expansion's Education and Labor Market Effects Using Two Before-expansion Survey Rounds (i.e., CHIP 1988 & 1995). 69

Table 2.6: Falsification Tests - DiD Estimators of the Expansion's Education and Labor Market Effects Using Two Before-expansion College Cohorts (i.e., 1985-91 vs. 1992-98). ..70

Table 2.7: Robustness Checks – Excluding Beijing and Shanghai, with and without Province Covariates. 75

Table 3.1: Age-group Specific Relative Supply Index for BA- vs. HS-educated Labor by Year. 99

Table 3.2: Effects of Supply, Demand, and Age on BA-HS Earnings Gaps. 105

Acknowledgments

I would like to express my deepest and most sincere gratitude to my advisor and dissertation sponsor, Prof. Judith Scott-Clayton, for her continued guidance, support, and encouragement during my Ph.D. journey.

This work would not have been accompanied without the support and help from my dissertation committee members: Prof. Sarah Cohodes, Prof. Judith Scott-Clayton, Prof. Alex Eble, Prof. Aaron Pallas, and Prof. Qin Gao, especially during an unusual and challenging time facing the COVID-19 pandemic.

I also benefitted tremendously from inspiring conversations with Prof. Henry Levin, Prof. Mun Tsang, and Prof. Thomas Bailey, as well as with professors in China who attended my job talks at Peking University and Beijing Normal University and gave me insightful feedback.

I am deeply grateful for the National Academy of Education/Spencer Foundation, Columbia University's Weatherhead East Asian Institute, and the Department of Education Policy and Social Analysis at Teachers College, for awarding me three dissertation awards that both financially support my study and provide amazing mentorship opportunities.

Lastly, words cannot express my gratitude for my parents, Lijun Wen and Shaohua Qiao, for their unconditional love and support. They are the greatest parents anyone can ever have, and I am blessed to have born and raised in the way I am.

Dedication

This dissertation is dedicated to my parents, extended family, and

all my teachers, professors, and mentors since kindergarten

for their love, encouragement, expectation, and inspiration.

Chapter 1: Estimating Labor Market Effects of Large-scale Education Programs: A Review of the Literature with Application to China's Higher Education Expansion

1.1 Introduction

Prior literature estimating the relationship between education and labor market outcomes is extensive. The pioneering studies in this field focus on using one cross-section of individuals to study the relationship between wage and its determinants, including years of schooling (Becker, 1962, 1994; Becker & Chiswick, 1966; Griliches, 1977; Mincer, 1974). For example, the canonical Mincer earnings model regresses the natural logarithm of earnings on a linear term of years of schooling, linear and quadratic terms of years of working experience, and a constant. The derived coefficient on one additional year of schooling, or the rate of return to education, could provide information to guide individual decision on human capital investment, facilitate efficient allocation of between physical and human capital, and have important implications for income distribution.

One caveat of the Mincerian model is that if schooling is endogenous, the coefficient on "education" using a simple ordinary least squares (OLS) function will be biased. More specifically, if individuals self-select into different educational levels based on observed or unobserved factors that potentially also affect their future earnings, the estimated effect of education on earnings will be biased. Even after controlling for observed variables, any omitted variables in the error term or variables that measured with error could still bias the coefficient on schooling. For example, if we cannot fully control for ability, which is positively related to both education and earnings, the resulting estimate of the returns to education omitting ability will be biased upward. Besides, if "education" is measured with lots of noise, the resulting estimate of the returns to education with substantial measurement error will be biased towards zero.

This issue of the endogeneity of education inspires a large body of later work that aims to find credible exogenous variation in schooling and thus to estimate the "causal" effect of education on earnings. These studies usually employ treatment effect models based on experiments or quasi-experimental methods. Among them, micro interventions that use randomized controlled trials (RCTs) to allocate some aspects of "education" are the most clean-cut method to identify the treatment or program effects. Other quasi-experimental methods used to obtain a causal interpretation of education's impacts include instrumental variable (IV), regression discontinuity (RD), and difference-in-differences (DiD) (Hoekstra, 2009; Ost, Pan, & Webber, 2018; Zimmerman, 2014), matching (Brand & Xie, 2010), fixed-effect strategies (Jepsen, Troske, & Coomes, 2014), synthetic control method (Abadie, Diamond, & Hainmueller, 2015; Ridley & Terrier, 2018), and the marginal treatment effect framework (Brinch, Mogstad, & Wiswall, 2017; Heckman, Lochner, & Todd, 2006; Kline & Walters, 2016).[1]

This "treatment effect" studies are generally within the "partial equilibrium" framework (Heckman & Robb, 1985). They intend to identify the causal effect of a specific "treatment," holding all other conditions to be fixed. This approach works best in micro-interventions when the effects of the intervention on the market or economy at large are ignorable. However, when

---

[1] For more detailed reviews of prior literature on returns to schooling, see Card (1995, 1999); Heckman, Lochner, & Todd (2003, 2006, 2008); Oreopoulos & Salvanes (2011); and Oreopoulos & Petronijevic (2013).

the intervention or policy under evaluation is large-scale in nature, and thus induces "general equilibrium" effects (for example, the equilibrium skill price of college-educated labor is likely to be affected due to China's college expansion program that translates into increases in the supply of college-educated labor), the assumptions underlying most of the experimental and quasi-experimental methods will not hold. For instance, the Rubin causal framework of potential outcomes (1978) has the Stable Unit Treatment Value Assumption (SUTVA), which assumes no spillovers to those who are not treated, either directly from treated units to untreated units, or through general equilibrium mechanisms (like skill prices change) that would also affect the non-participants. Consequently, evaluations of large-scale programs or policy changes using partial equilibrium treatment effect models might lead to misleading results. Moreover, the estimated effects of small-scale programs, evenly credibly identified, might not be able to predict what would happen when such programs are scaled up and affect the economy at large.

In addition to these assumptions, most partial equilibrium treatment effect studies rely on reduced-form models to evaluate programs or policy changes. They are straightforward and involve the least assumptions, but at the same time are frequently criticized of being lack of external validity, not relying on economic theory, and being unable to uncover mechanisms (Low & Meghir, 2017).

Building upon these limitations, recent work of policy evaluation incorporates general equilibrium analyses that relax assumptions under the partial equilibrium condition (Heckman, Lochner, & Tabler, 1998a, 1998b; Fernandez & Rogerson, 1998; Lee & Wolpin, 2006). General equilibrium analyses usually employ static or dynamic structural models that allow for relative skill prices to endogenously adjust to shifts in the supply or demand of different kinds of labor over time. Besides, they can simulate program effects beyond the specific context and samples from where their parameters are derived from. Nevertheless, these advantages do not come for free since general equilibrium models involve more assumptions during model set-up, and require more complex data to specify the model.[2]

The reminder of this review is organized as follows. In Section 2, I briefly summarize the treatment effect literature, and use three empirical papers to demonstrate how quasi-experimental methods could be applied to study education expansion programs that are most relevant to the Chinese case. I also explain how the partial-equilibrium returns to education estimated using micro-interventions differ from the aggregate effects of large-scale education programs, and why the reduced-form models can no longer separate the partial-equilibrium returns to education from other spillover effects. In Section 3, I shift to the general equilibrium literature and focus on papers that build models of skill price determination to explain the evolution of wage structure in the labor market over time. In Section 4, I use two recent studies to illustrate that these two frameworks could be integrated regardless of their seemingly distinct methodological approaches. Section 5 provides background information on the Chinese context, including a brief introduction to China's higher education system, the college expansion program implemented in 1999, prior literature on the returns to education in China, and previous studies evaluating the effects of the college expansion. Section 6 summarizes.

1.2 Using Large-scale Education Programs as Quasi-experiments

The treatment effect literature intends to identify an intervention or policy's causal impact on a specific outcome, such as years of schooling and earnings, and generally assumes

[2] Other directions that refine the basic framework and original model include: modelling risks associated with making educational choices and costs of schooling (Heckman, Lochner, & Todd, 2006, 2008), estimating distributional effects of the educational returns (Heckman & Vytlacil, 2005), and studying externalities in addition to private monetary returns associated with education (Acemoglu & Angrist, 2001; Oreopoulos & Salvanes, 2011).

that all other conditions are held constant as individuals get "treated." This literature is extensive and long-standing, and a full review is beyond the scope of this paper (for recent detailed reviews of various experimental and quasi-experimental methods frequently used by treatment effect models, see Angrist & Krueger, 2011; Angrist & Pischke, 2014; Athey & Imbens, 2017; DiNardo & Lee, 2011; Imbens & Rubin, 2015; Imbens & Wooldridge, 2009; Murnane & Willett, 2011; and VanderWeele, 2015; among many others. For applications specifically to education, see Schlotter, Schwerdt, Woessnann, 2009; Schneider, Carnoy, Kilpatrick, Schmidt, & Shavelson, 2007). In general, the quasi-experimental estimates of the effects of education obtained from these studies are usually larger than those estimated using simple OLS regression models, partly because these local average treatment effect (LATE) estimates applies to a "marginal" sample of individuals whose education decisions are induced to change by the treatment, and these individuals could have higher returns to education than the population average (for discussion about the "local" nature of quasi-experimental estimates, see Imbens & Angrist, 1994; Heckman & Vytlacil, 2005).

Rather than reviewing the entire literature, below I review three influential papers in the treatment effect literature that are most relevant to my proposed analysis of China's college expansion program. These three studies examine two national education expansion programs that share three features with the Chinese one: 1) they are all designed to raise educational attainment, 2) the programs are large-scale as a whole, and have variation in program intensity across regions, and 3) the three analyses focus on examining the programs' effects on labor market outcomes.

Lemieux and Card (2001) use the Canadian G.I. Bill to study the causal effects of college attendance on earnings. Like the G.I Bill in the U.S., Canadian veterans of the second World

War benefited from an educational program called Veteran's Rehabilitation Act (VRA). What is unique with Canada is that there were interprovincial variations in military enlistment rates (and thus the proportion of people eligible for the program) and in their education systems that made it easy or difficult for eligible people to attend university. More specifically, the authors study men from Ontario and Quebec, the two largest provinces in Canada that have similar industrial structures, and find that for those born in the early 1920s, over 60 percent of Ontario men served in the war, compared with a rate of 15 percent for Quebec men. Moreover, the elitist Quebec education system made it difficult for eligible veterans to get admittance to universities in Quebec. These two facts combined made the VRA program has roughly no effect on higher education attendance in Quebec, which makes it a credible control group.[3] The authors use this backdrop and thus adopt an IV approach to analyze the program's impacts on education and earnings for men from Ontario, using their counterparts from Quebec as a control group.

Using data from the 1971 Census and use potential eligibility for VRA benefits (a dummy indicating whether born in Ontario and born in the early 1920s) as an IV, they find that the program raised education for Ontario men born in the early 1920s by 0.2 to 0.4 years. Note that their sample includes all men in a given cohort and province, not just veterans who might be a selective sub-sample of the larger population. Their IV estimate of the rate of return to education is 15 percent, which is larger than the corresponding OLS estimates of 7 percent.

As an alternative specification, the authors also use the interaction between veteran status and family background as an IV for education, under the assumption that veterans from

[3] The authors argue that the large-scale G.I. bill in the U.S. suffers from the lack of a credible control group: first, there was no variation across regions in terms of program eligibility; and second, even after the war, a substantial number of people serve in the military and were eligible for the program, making the post-war cohorts unable to serve as an ideal control group either.

disadvantaged family background are more likely to take advantage of the program. In doing so, they allow being a veteran to have a direct effect on earnings. Using data from the 1973 Canadian Job Mobility Survey (in which there is information on veteran status that is absent in the Census), they find that the direct effect of WWII service on earnings is negative, implying that the IV estimates using the Census (that allows the war service to affect earnings only through education) could understate the true returns to education.

The authors also note that their IV estimates could be contaminated by potential spill-over effects due to the surge in the supply of college-educated individuals of specific cohorts (i.e., Ontario men born in the 1920s). But, the specific magnitude of such effects is unknown, which depends on the elasticity of substitution across cohorts: if there's perfect substitutions across cohorts, the IV estimates cannot capture any spill-over effects because increases in cohort-specific college graduates affect college graduates in all cohorts; on the other hand, if there's no substitution across cohorts, the IV estimates capture the full spill-over effects because increase in the supply of college graduates in a given cohort only depresses the skill price for this specific cohort.[4]

The other two papers by Duflo (2001, 2004) study a large-scale primary school construction program in Indonesia. During 1973-1978, the Indonesian government built more than 61,000 primary schools, which is roughly 2 schools per 1,000 children aged 5-14 in 1971. Enrollment rates among 7-12-year-olds increased from 69 percent in 1973 to 83 percent in 1978. However, the number of new schools built varied across districts, with more schools built in areas with lower primary-school enrollment. The first paper by the author (Duflo, 2001) takes

[4] Card and Lemieux (2001) find that the elasticity of substitution across cohorts is 5, suggesting that workers of different cohorts are strong but not perfect substitutes.

advantage of this fact and thus adopts DiD and IV strategies to estimate the program's impacts on education and earnings.

Using a cross-section of men born between 1950-1972 from the 1995 inter-censal survey of Indonesia (SUPAS), the author uses the interaction between the intensity of the program in an individual's region of birth and a dummy indicting whether the individual belongs to the young cohorts that would be affected by the school building program as an IV for years of schooling. For program intensity, the author uses two strategies: she first divides districts into "high program intensity" and "low program intensity" groups, and uses a DiD strategy. She also uses a continuous district-level program intensity measure, calculated as the number of new schools built out of the total number of primary-school-age children in an individual's region of birth. Moreover, besides using a dummy indicting whether an individual would be affected by the school building program, the author also uses exact cohort years (to interact with program intensity), thus coefficients represent the program's impacts on specific cohorts and can therefore be used to show cohort-by-cohort contrasts. For covariates, the author controls for cohort of birth and district of birth fixed effects. She also controls for the interactions between the "younger" dummy and a vector of time-varying region-specific variables that could be correlated with the program. For example, the author includes region of birth enrollment rate in the pre-program level and whether the district received another large development program happened concurrently during the primary school construction program under study.

Findings show that for cohorts fully exposed to the program, one primary school constructed per 1,000 children increased schooling by 0.12-0.19 years, and earnings by 1.5-2.7 percent, suggesting a 2-stage least squares (2SLS) estimate of the rate of return to education of 6.8-10.6 percent. The author also conducts subgroup analysis by pre-program population density, poverty rate and educational level. She finds that program effects are larger in sparsely populated regions, which implies that the increases might work through the channel that newly-built schools reduce the distance to school. Moreover, the author verifies that changes in the average pupil-teacher ratio during the program period are not systematically related to the number of new schools constructed in each region, suggesting that changes in the quality of education do not drive the estimated program effects.

One caveat of this analysis is that given the magnitude of the school-building program, there could be spillover effects for the untreated individuals, that is, who were not in primary school age during the program period or in areas where schools were built. Therefore, in a follow-up study, Duflo (2004) extends her previous analysis to take into account one aspect of the general equilibrium effects of the same program, namely, its effects on older cohorts who did not benefit from the school building program. Due to the program, the proportion of primary-school-educated labor increased since 1986 (the year the first affected cohort entered the labor market). Using panel data covering the years of 1986-1999, and using the interactions between the number of new schools built per 1,000 children and the survey year as instruments for the fraction of primary-school-educated workers in the region, she finds that earnings for older cohorts who were not exposed to the program fall as the share of primary school graduates rises: an increase of 10 percentage points in the proportion of primary school graduates in the labor force depresses older cohorts' wage by 3.8-10 percent. Apart from average wage for primary school graduates, the wage gap between primary school graduates and dropouts for untreated older workers falls as well. These results hold after controlling for baseline characteristics that determine the initial allocation of new schools, and may therefore have caused different growth trajectories across regions. The findings confirm that large-scale educational programs could have spill-over effects on the sub-population who were not directly affected by them. This again underscores the importance of considering the general equilibrium effects when evaluating large-scale programs and policy changes.

All the three papers show that the programs do increase students' educational attainments as intended, but their impacts on treated students' labor market outcomes are more complex. Moreover, these large-scale programs also have spillover effects to untreated individuals in the labor market. As a result, the quasi-experimental estimates of the earnings effects of the programs are likely to contain both the returns to increased schooling and other spillover effects caused by the program itself. Interpreting these aggregate or net earnings effects as the partial equilibrium "causal" returns to schooling could be misleading.

1.3 Structural Models of Labor Market Dynamics

The previous section demonstrates that large-scale policy changes tend to alter labor market dynamics and affect people who are not directly "treated," leading to general equilibrium effects.

Acemoglu (2010) identifies three types of general equilibrium effect that are usually left out in partial equilibrium effect analysis. First, factor productivities and prices could change (due to diminishing returns and imperfect substitution between factors) in respond to large-scale interventions or shocks. For instance, in the context of substantial college expansion, the surge in the supply of college-educated labor could depress the equilibrium wage for college-educated labor, and thus dampens the returns to college for all individuals with a college degree. Second, there could be compositional effects. The quality composition of the high school- and college-educated labor could change after the expansion, either because lower-ability students enter college, or the average quality of college education declines. Both effects are likely to offset or even reverse the partial equilibrium effects for small-scale interventions. Third, the policy changes or shocks could lead to endogenous technology responses: this type of general equilibrium effect could either attenuate or enhance the partial equilibrium effect. For example, Acemoglu (1998) shows that the increasing returns to the college premium in late 1970s in the U.S. under the increasing relative supply of college-educated workers since the late 1960s are the consequence of the endogenous response of technology to the relative profusion of high-skill workers.[5]

A couple of studies directly compare an intervention's effects in partial- and general-equilibriums. For example, Svensson and Yanagizawa-Drott (2012) evaluate the effects of an intervention in rural Uganda that distributes information on urban market prices of food crops via radio stations. The authors estimate the intervention's effects under partial and general equilibrium assumptions respectively, and reach very different conclusions. In partial equilibrium that assumes fixed market price and no spillovers to untreated farmers, the intervention has a significant and positive impact (40-55%) on crop revenue for farmers who have radio that broadcasts crops' market price. However, when millions of farmers receive the broadcasting and thus respond by increasing their supply of crops, the intervention has a general equilibrium effect on the crop price in the urban market. Under the latter scenario, the intervention has negligible aggregate effects (-2%) on crop revenue, as well as large distributional consequences: farmers who do not have radio are worse off because they respond to lower crop price due to increased supply from informed farmers by reducing their supply.

[5] For large-scale interventions, there may also be a "seesaw" effect (Acemoglu, Johnson, Robinson, & Querubin, 2008) on top of general equilibrium effects, which means the effect of an intervention that benefits some groups of the population could be undermined by other powerful groups who want to maintain their interest (i.e., the intervention could lead to endogenous policy responses that weaken its designed objectives).

Particularly with returns to (college) education, Heckman, Lochner and Taber (1998a) evaluate a national tuition reduction policy that induces a substantial increase in college enrollment. They find that estimated returns to college enrollment are more than ten times larger in the partial equilibrium framework than the long-run general equilibrium effects. Another study by Angrist (1995) shows that Palestinian students who attended college due to a rapid college expansion program in the West Bank and Gaza Strip earn less than half of the college premiums in general equilibrium (less than 20 percent) than those would be in partial equilibrium (40 percent).

These discrepancies underscore the importance of differentiating the returns to education from the aggregate program effect when studying large-scale educational programs. Under a basic supply and demand model, one could easily conjecture that educational expansions would increase the supply of high-skill labor, which in turn would depress their skill price. This is under the assumptions that: (1) perfect substitution across cohorts or age groups with the same level of skill, (2) relative demand for skill remains constant, and (3) the quality composition of two types of labor does not change. In the following paragraphs, I review papers that relax these assumptions and thus enrich the basic supply-demand framework to study labor market dynamics. More specifically, Card and Lemieux (2001) relax the assumption of perfect substitution across cohorts to allow cohort-specific relative supply to directly affect wages for specific cohorts on top of the effects from shifts in aggregate supply that affect all cohorts. The other four papers further allow the demand curve for skill to change over time, and develop an emerging field of theories on the relation between technological change and wage dynamics. Lastly, Carneiro and Lee (2011) incorporate measures of labor quality into their model, allowing the composition of each type of workers to change over time, which has a direct effect on wage dynamics in addition to supply and demand factors. I will elaborate more on each of these papers in the following paragraphs.

1.3.1 Shifts of the aggregate and cohort-specific supply curves

Card and Lemieux (2001) find that the evolution of college premium (or the wage gap between high school and college graduates) follow differential trends for male workers of different birth cohorts during the 1970s-1990s in the United States. Young cohorts enjoyed a rising college premium, while that for older workers stagnated during the same time. They also observe similar patterns for workers in the UK and Canada. The basic supply-demand model, which posits that the relative *aggregate* supply of two types of workers determine their wage gap, cannot explain these un-parallel trends for workers from different birth cohorts. Therefore, the authors extend the basic model to allow for imperfect substitution between workers of different age groups and with the same education level, and thus allow *cohort-specific* relative supply of the college-educated labor to directly affect the college premium on top of the effect due to aggregate relative supply of these two types of labor.

The authors use data from the U.S., UK, and Canada to specify the parameters in their model of labor supply with imperfect substitution between age groups. They find that the observed distinct patterns of a rising college premium for the baby-boom cohorts in the three countries are mainly due to a slow-down of the growth in educational attainment for these cohorts, which later transfers into a decrease in the cohort-specific relative supply of college-educated labor for these cohorts. Besides these cohorts, their models for the three countries uniformly suggest that the evolution of college premium during the 1970s-1990s can be largely explained by: (1) cohort-specific relative supply of college-educated labor, and (2) steady rises in the relative productivity of college-educated workers (due to skill-biased technical change) that translate to a constant rate of growth in the aggregate relative demand for college-educated labor. The implications of their findings suggest that cohorts who experienced substantial changes in the cohort-wide average educational attainment either due to demographic reasons or policy changes could experience substantial cohort effects in addition to the effects equally spread out to all cohorts within each education level.

1.3.2 Shifts of the demand curve

Several studies attempt to model and explain shifts of the demand curves for different types of labor (on top of any changes in the supply curve) to explain changes in income inequality in the labor market. Some of these works are motivated by the observation that returns to skill, or the college premium that is usually measured as the wage gap between college- and HS-educated workers, have been increasing. Meanwhile, the relative supply of college-educated workers has been increasing as well, which suggests that there should be a concurrent outward shift of the relative demand curve for college-educated labor.

The pioneering work by Tinbergen (1974) links the relative demand for skill to technology, particularly to the skill-bias of technical change ("bias" in the sense that technology increases productivity of high-skill labor more than that of low-skill labor). Tinbergen models a relative technology efficiency parameter between the high- and low-skill workers ($\frac{A_H}{A_L}$) that represents the relative productivity between the two types of labor into the production function (in CES form) for the aggregate economy. The model shows that skill-biased technological growth will widen the wage gap (assuming that high- and low-skill workers are complements), while increases in the relative supply of high-skill workers will narrow the gap.

Goldin and Katz (2007a) empirically fit Tinbergen's model using data from the Current Population Survey (CPS) for years covering 1963 to 1987, a period when both the average educational levels and the relative demand for high-skill labor increase. Their results show that during 1963-1987, the estimated average annual increase in the relative demand for high-skill labor is 2.7 percent. Acemoglu and Autor (2010) extend their estimate using longer period of data during 1963-2008, and find that the estimated average annual growth in relative demand for high-skill labor becomes smaller, which is roughly 1.6 percent. This increasing relative demand for high-skill labor will raise the relative wage for high-skill labor, and thus widen the wage gap between the high- and low-skill labor.

The above-mentioned models illustrate that technological advancement boosts the demand for high-skill (or college-educated) labor, whereas human capital investment increases the supply of such labor. When demand for high-skill labor moves outward faster than does the supply of them, the wage gap between high- and low-skill labor (or college- and high school-educated labor) widens; and vice versa when supply outpaces demand. This is the so-called "the race between education and technology" (Goldin & Katz, 2009).

The study by Johnson (1997) tries to explain the driving forces for the shifts of the demand curve. The author concludes with two: trade and openness of the economy, and skill-biased technological change. The author emphasizes on the second reason, and categorizes technological changes into four types: (1) intensive skill-biased technological change, in which high-skill workers become more productive in jobs they formerly performed. For example, personal computers disproportionally raise the productivity of high-skill workers than that of low-skill workers; (2) extensive skill-biased technological change, in which high-skill workers become more productive in jobs that were formerly done by low-skill workers. The author uses the introduction of robotics to manufactory as an example to illustrate this type of technological changes since the adoption of robotics complicates many jobs that were formerly routine and required only low-skill workers. After the introduction of robotics, more engineers were demanded for positions formerly accomplished by low-skill labor; (3) skill-neutral technological change, in which technology advancement increases all skill groups' productivity proportionally; and (4) "unskill-bias" technology change that causes a shift from skilled toward unskilled labor. For instance, with the possibility of purchasing insurance policies online, some insurance sales have been replaced with high school graduates who can handle digitalized purchases. This last type of technology change is the opposite to the extensive skill-biased technological change.

Among the four types of technological changes, the first one will shift the relative demand curve to the right if the elasticity of substitution between high- and low-skill labor σ is greater than 1 (which is generally the case); the second type will shift the relative demand curve to the right regardless of the value of σ; the relative demand curve won't change for the third type of technological change; and the fourth type will shift the relative demand curve to the left. This classification put forth more than twenty years ago still has explanatory power today, and the different types of technological changes outlined in this work could still guide the modelling of the shifts in the demand curves.

All the previously mentioned models of technological advancement would predict that real wages of both types of labor would increase over time. However, Acemoglu and Autor (2010) note that there are discrepancies between what these models would predict and the actual data observed more recently, including falling real wages for some skill groups, polarization of income growth, and polarization of occupational growth.

They therefore introduce the concept of "task" to extend the previous skill-based theories and models. They distinguish tasks from skills in that "a skill is a worker's stock of capabilities for performing various tasks" (Acemoglu & Autor, 2010). The assignment of skills to tasks can evolve with time, either because shifts in market prices mandate reallocation of skills to tasks, or because the set of tasks demanded in economy is altered by technological development, trade, or offshoring. One noteworthy feature of this framework is that it allows for task-replacing technological change, which can directly reduce the real wages for certain workers who formerly performing those tasks, even if it would raise total output at the same time.

They apply their task-based framework to study the relationship between state-level per capita income growth and state-level relative supply of skill, technology (reflected in the range of tasks in use), and the interaction between education and technology during 1960-2007. One interesting finding from their analysis is that the growth in higher education contributes to larger increases in per capita income in states where the *initial* level of job complexity in the labor market is higher. In other words, the returns to education could depend on the conditions of the labor market, such as the set of tasks in use and the complexity of such tasks.

1.3.3 Changes in quality composition of different types of labor

The last study in this section incorporates changes in the quality composition of different types of labor into the model. Carneiro and Lee (2011) point out that studies on the evolution of college premium using panel data always assume that the relative quality composition of certain types of workers is constant over time. However, this might not be the case, especially when the changes in the relative supply were due to some radical and large-scale policies that altered the educational choices of different groups of individuals. Using data from 1960-2000, a period when there are gradual yet substantial increases in college attendance in the U.S., they decompose changes in the college premium into price effect that is determined by relative supply and demand of college-educated workers in an individual's region of work, and composition effect that captures changes in the average quality of individuals in each schooling level. For the latter, they measure cohort-specific relative quality of college-educated labor using the proportion of cohort-specific college enrollment out of college-age population in an individual's region of birth.

Results from their analysis show that from 1960 to 2000, the proportion of college-age population enrolled in college increased from about 30 percent to 60 percent. This increase led to a decline in the average quality of college graduates during this period, resulting in a 6 percentage points decrease in the college premium. Consequently, the composition adjusted college premium rose by 26 percent over 1960-2000, rather than the reduced-form college premium of 20 percent. This holds after controlling for state fixed effects and measures of quality of K-12 education at the state-year level, which could potentially affect both the proportion of college-goers and earnings.

Moreover, the decrease in quality could be due to either a reduction in quality at the margin because more marginal students entered college, or a reduction in quality on average due to resource crowding and/or peer effect. Although Carneiro and Lee's model cannot directly tease apart these factors, the authors find that the proportion of college-goers has negligible effects on the average high school wages. They interpret this finding as implying that change in the ability of marginal students (i.e., those who would be in the high school group if the proportion of college-goers did not increase) does not account for the major part of the quality decline, which in turn suggests that the cohort crowding hypothesis plays the major role in explaining this decline.

Broadly speaking, Carneiro and Lee's work has two major implications. First, leaving out composition effects could lead to an underestimate of the returns to college in the context of increasing relative supply of college-educated labor. Second, their model can directly estimate the effects of migrating from areas of birth to areas of work, which could have important policy implications. For example, in the context of a decentralization of school finance, if better-educated people could migrate to richer districts to work and thus earn a migration premium, poorer districts would have fewer incentives to invest in educating their local students. Bound, Kzedi, and Turner (2000) and Duflo (2004) discuss this issue in the context of both higher education and K-12 education, as well as developed and developing countries.

All types of general equilibrium effects outlined in this section could emerge during China's college expansion. Therefore, the models and methods discussed here can be applied to the evaluation of China's college expansion to achieve a fuller picture of its impacts on treated individuals and the labor market or economy at large.

1.4 Recent Studies Bridging the Two Approaches

From previous sections, we see that the relative advantage of partial equilibrium treatment effect analyses is its simplicity and the causal interpretation of the estimate on the treatment variable, whereas the relative advantage of general-equilibrium structural models lies in its theoretical foundation of the model parameters, its power to uncover the mechanisms through which effects operate and to predict policy counterfactuals and how a program could translate to other contexts. However, parameters in general-equilibrium structural models usually do not carry causal interpretations.

Some recent works bring together these two approaches by using structural models with exogenous sources of variation to identify treatment effects in general equilibrium (for example, see Kaboski & Townsend, 2011; Voena, 2015). Such studies demonstrate that combining experimental or quasi-experimental methods with structural models could bring together the best of both approaches: credible evaluations can be used either to validate a structural model or to aid in the estimation process; meanwhile structural model can add to the experimental evidence by uncovering mechanisms and allowing for analysis of counterfactuals (Low & Meghir, 2017). In the following paragraphs, I review two studies that take advantage of macro-level exogenous shocks from large-scale educational programs to specify their structural models, and thus obtain causal interpretations to at least some of their parameters.

Khanna (2015) combines the RD approach and a structural model of wage determination for two types of labor derived from the aggregate production function to study the general equilibrium effects of a national school construction program in India. During the 1990s and early 2000s, India initiated a school building program called "District Primary Education Program (DPEP)" that primarily built schools, hired teachers and upgraded infrastructure in low-literacy districts. The allocation rule of the schools is based on female literacy rate in 1991 and those below the national average are eligible for the program.

The author uses this program eligibility cutoff to identify his model of the wage gap between the high- and low-skill labor. More specifically, he combines three waves of a household survey, a census of firms, school-level data, test score surveys, and the Indian census that together assembles a 10-year panel. Then the model can be identified by separately estimating the program's effects on each skill-age group cell (i.e., skilled old, skilled young, unskilled old, unskilled young) on both sides of the RD cutoff, and thus parsing apart the returns to education in general equilibrium and other general equilibrium effects.

Findings show that the estimated returns to shifting into the skilled group in the absence of any general equilibrium effects are 19.9 percent (i.e., the partial equilibrium returns to education, holding all other situations constant); the effects of having one more year of education with general equilibrium effects are 13.4 percent (i.e., the aggregate earnings return to one additional year of schooling). The author also concludes that approximately 92 percent of the general equilibrium effects are borne by the young treated cohorts, whereas the two components in the general equilibrium effects for all cohorts (i.e., aggregate relative supply and demand for the skilled labor) counteract each other because both curves shifted outwards.

The other paper (Bianchi, 2015) estimates the general equilibrium effects of higher education expansion in Italy. In 1961, the Italian government lowered the requirement for students to enroll in university STEM majors, which were only open to students from university-prep high schools ("type A" students) before the policy change. Meanwhile, other university majors remain inaccessible to high school graduates from industry-prep high schools ("type B" students). By 1968, freshmen enrollment in STEM majors increased by 216 percent. The author intends to uncover how student learning in college and earnings are affected by the large inflow of potentially lower-ability students into the STEM majors.

The author links comprehensive transcripts data from both high school registries and universities at the "student-university course-academic year" level to a cross-section of earnings data in 2005, when the affected cohorts are 56-67 years old. Findings show that university enrollment and graduation of type B students increased; however, earnings did not increase for type B students. Meanwhile, substantial number of type A students shifted from STEM majors to other restricted majors. Moreover, course-level grades, used as a measure of human capital accumulation in college, decreased for all STEM major students.

He further decomposes the decline in returns to STEM majors into two parts, and finds that dropping human capital accounts for 42.9 percent, of which crowding of university resources accounts for 24.6 percent, while higher class heterogeneity that made university teaching less effective could explain the other 18.3 percent. The remaining part is captured by depressed skill prices for STEM majors due to increase in the supply of them.

This work contributes in the following two major ways. First, it further decomposes the effect of a decrease in quality into resources crowding and changes in student composition. This distinction has important policy implications, since the former one is easier to fix, but the latter one is much harder. Works from others find similar finding that student composition affects short- and long-run student outcomes (Duflo, Dupas & Kremer, 2011; Figlio & Page, 2002; Lavy, Paserman, & Schlosser, 2012; Jacob & Lefgren, 2004), which further implies that policies that change class composition could result in unintended consequences. Second, this paper quantifies how expansion affects student learning in the higher education context. Though this cannot be achieved without the availability of very comprehensive transcripts and earnings datasets.

To sum up, both the Indian school construction and Italian higher education expansion papers incorporate structural models to decompose the general equilibrium effects of large-scale education expansion programs into supply, demand, and quality composition factors, and thus reveal the mechanisms through which the expansions affect the labor market dynamics. Their models are also able to predict and simulate out-of-sample scenarios and counterfactuals that could have important policy implications beyond specific country contexts.

1.5 The Chinese Context

In this section, I illustrate how the theories and methods mentioned in previous sections could apply similarly or differently to the Chinese context. This section contains a brief introduction to China's higher education system and the college expansion program starting in

1999, and an overview of previous studies on the returns to education and the expansion's consequences in China.

1.5.1 Returns to education in modern China

Returns to one additional year of schooling in China were estimated as between 3-4 percent during 1980s and 1990s (Byron & Manaloto, 1990; Johnson & Chow, 1997; Liu, 1998), much lower than the world average rate of return to education of roughly 9 percent (Psacharopoulos & Patrinos, 2018). The low returns were largely due to China's centrally-planned economy and the principle of "equal distribution of income" under the Soviet influences held at that time (Gao & Smyth, 2015).[6]

Returns to education began to raise since the mid-1990s, and were around 10 percent estimated using OLS regressions during the 2000s (Ge & Yang, 2011; Meng, 2012; Zhang, Zhao, Park, & Song, 2005). For example, using data from the Urban Household Survey (UHS), Zhang et al. (2005) find that the rate of return to one more year of schooling rises from 4.0 percent in 1988 to 10.2 percent in 2001, and is mainly driven by increases in the college wage premium. Using data from the same survey, Meng (2012) updates the analysis to the year of 2009, and finds that returns to junior high school, senior high school, and college (compared with primary school only) continue to increase after 2001, and peak at 2006, with the most substantial increases for returns to college. After 2006, returns to all these three educational levels drop. The estimated returns in above-mentioned works are based on basic Mincer equations with gender and region dummies as controls.

[6] Many benefits of a "better" job were not reflected in higher wages, but in forms of free or discounted housing, food/commodity vouchers, etc. Even in recent years, non-monetary job benefits are still important, which include better insurance coverage and pension plan, whether the company/organization is willing to sponsor the change of its employees' "hukou," which in turn determines many benefits the employees and their offspring could enjoy.

There are fewer studies that employ causal methods, and those studies generally find ***lower*** returns than the OLS estimates. For example, comparing a sub-sample of twins with different levels of education attainment from five cities in China, Li, Liu, and Zhang (2012) estimate the rate of return to an additional year of schooling to be 2.7-3.8 percent, as compared with 8 percent using the same data source but using variation in education across the full sample. Giles, Park, and Wang (2015) use the Cultural Revolution in 1966-1968 – a period in which education at the high school and college levels were largely disrupted - as an IV, and estimate a rate of return to education of 8 percent using urban wage data in 2001; the corresponding OLS estimate is 9.6 percent. With respect to higher education, specifically, they find that the IV estimate of returns to college is 37.1 percent after controlling for city and cohort fixed effects, which is again lower than the OLS estimate of 49.8 percent. These findings are in contrast to previous studies conducted in the U.S. that generally found larger IV estimates than the corresponding OLS ones (for reasons explained in Section 2). Results from China could imply that the problem of ability bias when using OLS models in estimating education returns is a much more serious problem under the Chinese context.

1.5.2 Higher education and the higher education expansion in China

Since the founding of the People's Republic of China in 1949 up to early 1980s, China organized its higher education system based on the Soviet model: there were no private universities, and all public higher education institutions (HEIs) were under the jurisdiction of the central or provincial governments. Nearly all aspects of higher education, including admission, curricular, teaching, and job assignment of college graduates, were planned by the central government and organized using a top-down method.

In December of 1978, China implemented a national campion of "economic reform and opening-up," which transformed China from a planning economy to a market-oriented one. Shortly after China's economic reform, China's higher education system abandoned the Soviet influences, and HEIs gradually gained some autonomy. A series of reforms happened starting from the mid-1980s, including reorganization and merge of existing HEIs, diversification of higher education finance, and re-emergence of private HEIs. From 1993 to 2001, 708 small universities and colleges were re-organized into 302 larger institutions through the elimination of the line ministries' control over HEIs in China. Besides, HEIs were allowed to generate revenues through research contracts with industries, consulting services with business, training services, and fund-raising. Lastly, private HEIs in China re-emerged since 1980s to meet the increasing demand for higher education from the public. The majority of private HEIs in China are three-year community colleges that only grant associate's degrees, and they have been peripheral compared with public HEIs in the higher education market in China until the 2000s (Min, 2004), but are gaining increasing importance more recently.

Higher education admission system in China. China has a centralized higher education admission system: first, every year the central government sets a quota for the total number of students that could be admitted into regular HEIs nationwide, and second, admission into roughly all HEIs only depends on students' scores on the national college entrance exam (CEE).[7] At the same time, administration of the CEE and distribution of quota are operated at the provincial level, and follow the principal of "distributing quotas by (students' household registration) province and admitting students by their CEE scores" (Ministry of Education of

[7] The CEE is a pre-requisite for entering almost all HEIs in China, and is usually taken at the summer after students' senior high school.

People's Republic of China [MoE], 1996).[8] In practice, every year before the CEE, each HEI decides how many students from each province could be admitted in that Fall. Then, the institution-specific quotas are aggregated at students' province level, which would roughly determine how many students from specific provinces could attend any HEIs in China in that year. Consequently, quotas are pre-determined before students taking the CEE, and thus won't be affected by students' performance on the test.

How the exact number of quota is determined and distributed across provinces and HEIs each year is quite the black box. HEIs tend to allocate more spots to the province where they locate, though this can differ depending on the administrative feature of the HEIs. HEIs under provincial jurisdiction allocate higher proportion of spots to local students than institutions under central government or ministry control. Moreover, Zhou (2017) finds a positive correlation between province-level average CEE scores and the increase of provincial quota for next year, suggesting that HEIs adjust the quotas based on previous students' performance on the CEE.

On the students' side, students take the CEE and rank against each other based on their CEE scores within their province of household registration (or "hukou"). Students in all provinces take the exam at the same time (in early June each year), but provinces can administer their own exam questions. In theory, a student's "hukou" province and thus the number of admission quota assigned to his/her province, score on the CEE, and choices listed on the college application form will determine whether he/she could attend any college at all, and which institution he/she can attend. In 2006, 57 percent of students who took the CEE were admitted into at least one HEIs (MoE, 2006), and among those who attended college, 65 percent attended a HEI in their home province (Loyalka, 2009). As for students who are not admitted into any

[8] "Province" refers to students' birth province, not the province where HEIs locate.

HEIs or not satisfied with their admission results, they can re-take the CEE in next year(s).[9] A non-negligible number of students re-take the exam: according to data from one province in China, 37 percent of the freshmen students enrolled in all the 4-year universities in that province in 2008 have taken the CEE for more than once (Loyalka, 2009).

Once students are admitted into a university and a major, they rarely have the chance to transfer to other majors, and transferring between universities is even harder and rare.[10] On the other hand, once students are admitted into colleges, very few fail to complete their degrees unless they voluntarily choose to quit. For example, the average dropout rate in all 4-year universities in one prove in China for the 2008 cohort was less than 1 percent (Loyalka, 2009).

Types and tiers of HEIs in China, sources of funding, and returns to institutional quality. HEIs in China can be roughly categorized into four tiers. The first-tier consists of public universities under central government and the most selective flagship universities under provincial government; the second-tier includes less selective public universities; the third-tier includes private HEIs, and independent schools that are private subsidiaries of the 1st- and 2nd-tier universities; and the fourth-tier are three-year vocational schools, which are predominately private. In 2016, among all 2,596 HEIs, 1,855 (71.5%) are public: of which the majority (1,737 HEIs) are affiliated with provincial governments, and the other 118 are affiliated with the central government or ministries.

There are three main sources of revenue for HEIs: appropriations from the central and local governments, tuition and fees, and other sources like revenues from contracts with

[9] Alternatively, they could also choose to enter the labor force, or attend private community colleges that do not require CEE scores.

[10] Recently, more universities start to allow students to change majors or choose majors after the first year.

industries. Appropriations from the central government have been concentrated in the first-tier HEIs, and appropriations from local governments have been disproportionally provided to the first two tiers of HEIs. The third- and fourth-tier of HEIs seldom receive support from the governments. In terms of tuition and fees, higher education was free during 1949-1988 in China, and in 1989 tuition fees of approximately $30 per year was charged at two public universities. In 1997, tuition and fees were rolled out to the entire country, and China thus shifted from a free higher education system to a cost-sharing scheme in which listed tuition and fees are regulated (annually) to a level that are approximately the same across all the HEIs nationwide (MoE, 2020a). In recent years, the tuition and fees have been kept at approximately $1,000 dollars per year.

The school quality literature often uses measures such as per-student-expenditure and/or the average test score of incoming students to measure institutional quality and selectivity (Altonjo & Zimmerman, 2017; Hoxby, 2017; Minaya & Scott-Clayton, 2016), and usually finds positive (causal) effects of school quality on students' educational and labor market outcomes, at least for certain groups of students (Black & Smith, 2004; Brand & Halaby, 2006; Dale & Krueger, 2002, 2011; Hoekstra, 2009; Long, 2008). Under China's centralized and highly hierarchical higher education system, the positive correlations among student ability, institutional resource, and education quality are strong. However, there are few causal studies on the effects of college quality on student outcomes within the Chinese context, largely due to data limitations. For instance, CEE scores are hard to get, which, if available, would be ideal for a RD analysis; institutional-level resource measures are only available until recently (Yu, 2017). Two recent studies that do obtain such data find positive effects of attending elite colleges on graduates' labor market outcomes, including lower unemployment probability, higher wages, and higher esteem of the industry and occupation graduates work for (Jia & Li, 2016; Yu, 2017). Using a RD design, Fan, Meng, Wei, and Zhao (2010) find the causal returns to attending four-year universities (compared with being admitted into a three-year college) to be in the range of 40-60 percent.

Rationales for higher education expansion in China. In 1999, the Chinese government initiated a radical and large-scale higher education expansion program. In the year of 1999 alone, approximately 500,000 more students were admitted into China's regular higher education institutions, a 48 percent increase from 1998 (Educational Statistics Yearbook of China [ESYC], 1999). Between 1998 and 2001, nationwide enrollment in higher education more than doubled (ESYC, 2001).

Several factors have contributed to the expansion. Returns to college have been increasing during the late 1990s (see section 5.1 for more details), and China also enjoyed substantial Gross Domestic Product (GDP) growth at an annual rate of around 8 percent during the same time, resulting in rapid growth of household wealth. As suggested by the consumption theory (Fishlow, 1966), demand for a normal good would increase as real income rises. Both factors would stimulate the demand for higher education.[11] However, up to the year before the higher education expansion in 1998, there has been insufficient supply of higher education in China: gross enrollment rate for higher education in China, calculated as the number of full-time-equivalent students enrolling in HEIs out of the total number of 18-22 years old population, was constantly below 7 percent. Therefore, there was sizeable unmet needs for higher education during the pre-expansion era.

[11] In addition to the micro-level relations between college education and earnings, the Chinese government believes that education could promote economic growth, and human capital is important in global competition.

Except for this general background, two particular events also contributed to the higher education expansion. One is the reform for state-owned enterprises (SOE) starting in 1992, which led to considerable layoffs of workers in such organizations, driving up the unemployment rate during the late 1990s. The other one is the 1997 Asian financial crisis during which China's export was adversely affected since China insisted to keep its currency's value. Therefore, expanding the access to college seems a good remedy because first, it could keep more students in school for three or four years and thus postpone the potential unemployment issues already serious due to the SOE reform. Moreover, getting more students into college could stimulate domestic consumption and investment from school construction, which could help offset the negative impacts of shrinking export caused by the financial crisis.[12]

Lanqing Li, Vice Premier in charge of education at that time, recalled five reasons for the expansion in his book (Li, 2003): (1) letting more students go to college could help meet the need for more talented personnel to sustain the rapid development of the Chinese economy; (2) the public demand for higher education is increasing, and the government has the obligation to meet their demand; (3) college expansion could postpone the employment of high school graduates; (4) expanding higher education could increase education-related consumption and investment to stimulate domestic consumption and promote investment in related industries, such as construction and services; and (5) expanding college admission quota could alleviate the pressure on high school students and high schools, discouraging test-oriented teaching and learning, and thus promote all-around education in primary and secondary schools. Ms. Zhili

[12] Min Tang, an economist at the Asian Development Bank at that time who was said to have propose the expansion plan to China's Premier, estimated that if higher education enrollment could be doubled in three years, domestic consumption would be raised by about 100 billion yuan (approximately 12.5 billion U.S. dollars using exchange rate at that time), which amounts to 0.5 percent of the GDP (Min, 2004).

Chen, the Minister of Education at that time, also promoted a "non-conventional" expansion of higher education to reduce the gap in college attainment between China and the developed countries, and thus to promote China's competitive advantage in the global labor market in the 21st centaury (Huang, 2001).

As a result, in 1999, a plan to increase enrollment by 47 percent in that single year, and to achieve the threshold of a 15-percent gross enrollment rate by the year of 2005 was put forth.[13] It is worth emphasizing that the expansion was designed primarily under economic considerations, and even received some protests from the education sector. This creates some randomness to the event in the sense that schools, students, and leaders in the higher education sector did not expect such changes in advance.

1.5.3 Empirical studies on the causal effects of the higher education expansion on labor market outcomes

Previous studies that use the quasi-experimental nature of the policy change have examine the expansion's effects on the probability of employment and/or earnings. For example, using the 2000 and 2005 China Census data, Li, Whalley, and Xing (2014) are among the first to estimate how the expansion affects college graduates' probability of employment. Using a DiD method, they compare the changes in employment for young (22-24 years old) and older college graduates from 2000 to 2005. They found that unemployment rates increased by 6-9 percent for young college graduates in the 2005 Census data, who are likely to have attended college after the expansion. They thus conclude that the expansion decreases young college graduates'

[13] Trow (1972)'s theory of the three stages of higher education development was popular at that time in China, which classifies the higher education system as an elite one if GER was lower than 15 percent, and a mass one if the GER is between 15 to 50 percent; universal higher education is achieved if the GER is above 50 percent.

employment probability. Xing, Yang, and Li (2018) employ the same method and update the previous study by adding Census data from 2010. They reach similar conclusion: the expansion has negative impacts on young college graduates' employment probability, and the magnitude of the negative impacts is smaller when using the 2010 Census data as compared with Census 2005. One caveat of these two studies is that they restrict their analyses to college graduates only. However, part of the (or the major) educational effects of the expansion is to induce some high school graduates who would formerly directly enter the labor market to attend college. Their analyses cannot say anything about the expansion's employment impacts on these "treated" individuals. Moreover, the ability composition of college graduates could change substantially after the expansion.

Apart from employment, there are studies that explore the expansion's earning effects. This line of research usually use whether one is scheduled to attend college after the expansion as an IV for either college attainment or years of completed schooling to estimate the causal returns to education (see Hu, 2018; Wu & Zhao, 2010; for most recent works). The major problem with this method is that using only one dimension of the exogenous variation in college attainment - the timing of college attendance inferred by cohort of birth - to identify the treatment effects might suffer other confounding influences. For example, comparing two groups of 25-year-olds using earnings data from 2000 and 2010 could capture the earnings differentials between those unaffected and affected by the expansion. However, this earnings difference would also capture any macro-economic trends that happened concurrently during the expansion era that potentially affect the before- and after-expansion cohorts differently. Similarly, simply comparing earnings for the before- and after-expansion cohorts using one cross-section of earnings data cannot eliminate confounding influences from cohort effects.

More sophisticated work adds more dimension(s) of variation to control for confounding influences. For example, also using data from the 2005 Census, Ou and Zhao (2016) compare labor market outcomes for young exposed and old unexposed cohorts from provinces that experience different degrees of changes in college admission rates from 1998 to 2001. They find that the expansion lowered college graduates' probability of being unemployed, especially for male college graduates. This is in contrast to findings by Li et al. (2014) and Xing et al. (2018) that use the same dataset(s). Despite the decrease in unemployment probability, the authors also find that log hourly earnings for college graduates decrease, and the decrease is most substantial for high-skill and white-collar occupations. The method used by Ou and Zhao is more credible, though there are some limitations: first, the 2005 Census data could only capture the very early labor market outcomes for the first two or three cohorts that were affected by the 1999 expansion; second, the authors use the province level increase in the total number of students admitted into any colleges from 1998 to 2001 divided by the number of high school graduates in 1999 to measure the intensity of expansion among different provinces. Both the actual number of college admission and the number of high school graduates could be endogenous. For example, more students might decide to attend or persist in high school due to increased prospects of college attendance after the expansion. The number of planned college admission quota and college-aged population could be a better choice since they are less likely to be affected by the expansion. Last, like many other previous work on estimating the expansion's labor market consequences, Ou and Zhao also separately estimate earnings for college and high school graduates, and focus most of their analysis on college graduates only. As mentioned previously, this practice will lose the expansion's education effects (i.e., improve the average educational level of the labor force) and cannot answer whether they translate to better or worse labor market outcomes. It is thus interesting to see how the expansion affects the average level of earning for high school and college graduates combined.

Another group of studies does not intend to use any causal identification, but simply to depict changes in the returns to college after the expansion. These studies generally use panel or multiple years of cross-sectional data, and OLS or fixed-effect models. Returns to college have been increasing before the expansion, and are approximately 40 percent before 2002 when no post-expansion cohort has entered the labor market (Zhang et al., 2005). Researchers do not reach a consensus in terms of the trends in college premium afterwards: some studies show that college premium decline during the 2000s (Appleton, Song, & Xia, 2014; Yao, Xin, & Zhang, 2013), while others find the premium to kept increasing (Carnoy, Loyalka, Androushchak, & Proudnikova, 2012; Hu & Hibel, 2014; Yang & Gao, 2018) or stay relatively constant (Gao & Smyth, 2015) during the 2000s.

Among this line of research, one interesting study by Li, Liang, and Wu (2016) aims to depict and explain the changing wage structure in the labor market during the 1990s and 2000s. More specifically, they use repeated cross-sections of the annual Urban Household Survey during 1990-2009 to separately estimate college premium for two age groups: inexperienced (21-25 years old) and experienced (26-40 years old), controlling for gender and province of work fixed effects. They find that after the expansion, the college premium for inexperienced workers declined, but that for experienced workers increased during the same period. They also find that the experience premium - defined as the wage gap between experienced and inexperienced workers within the same level of education - increased dramatically after the expansion. They also build a structural model that allows for the supply of young inexperienced high-skill workers to endogenously respond to demand shocks, and conduct several simulations to test different hypotheses. They conclude that a demand shock for high-skill workers, such as a sudden inflow of foreign direct investment is the driving force for the pattern of changing wage structure observed for the past decades.

Another paper by Che and Zhang (2018) uses China's higher education expansion as an exogenous shock for the supply of high-skill labor to study its impacts on firm-level productivity growth. The authors use a DiD strategy and find that human-capital-intensive industries have a larger gain in total factor productivity (TFP) after 2003, that is when the first after-expansion cohort of college graduates starting to enter the labor market.[14] They also explore potential pathways and find that human-capital-intensive industries employ more high-skill workers, have higher research and development (R&D) expenditure, and also increase their physical capital adoption (reflected by importation of advanced capital goods) after 2003. They also find that foreign-owned firms have higher productivity gains than domestic private firms. Even though the stakeholders this work examines are not students or the labor market at large, findings for their analysis imply that the higher education expansion that raised the average level of human capital in the workforce has positive spillover effects on firm productivity, or more generally, on total output of the economy at large either due to technological advancement, and/or firms' more efficient utilization of and allocation between physical and human capital.

1.6 Summary

The higher education expansion program mandated by the Chinese government in 1999 relaxed the supply constraints of college education to some extent and substantially improved access to higher education among younger cohorts. However, previous studies sometimes find

[14] They measure human capital intensity of industry *j* by the percentage of workers with a four-year college education or more in industry *j* in the U.S. in 1980.

contradictory results regarding the expansion's labor market effects. Prior literature on the impacts of the expansion on employment and earnings seem to be very sensitive to the data and sample used, identification strategy employed, and measurement of key variables. This leaves room for more refined future research, especially to improve upon the following aspects: first, future work could use better data, such as more recent waves of earnings data that cover more after-expansion cohorts, and/or survey data that contain more comprehensive information on and more detailed and accurate measure of education; second, use more credible identification strategies and conduct falsification tests to validate identification assumptions, as well as robustness checks to test for the influences of potential cofounding factors; third, do not restrict the analyses to college graduates in order to evaluate the full impacts of the expansion; and last, pay attention to the spillover effects of the expansion, on top of its impacts on "treated" individuals, and explore mechanisms through which such effects take place. Theories, methods, and previous empirical work conducted in the context of other countries introduced in this chapter could help guide future analyses to help readers better understand one of the largest and most comprehensive educational reform on record and how conclusions may translate to other environments or circumstances.

Chapter 2: Estimating Causal Returns to College: Quasi-experimental Evidence from China's Higher Education Expansion

2.1 Introduction

The benefits of higher education are well-documented: college-educated individuals on average enjoy higher earnings and more stable employment (Heckman, Humphries, & Veramendi, 2018; Oreopoulos & Petronijevic, 2013), better health for one self and one's offspring (Currie & Moretti, 2003; Heckman, et. al, 2018; Mirowsky, 2017), less crime (Lochner & Moretti, 2004), and have higher level of civic engagement (Dee, 2004). In addition, a better-educated labor force promotes technological advancement and contributes to sustained economic growth (Aghion, Boustan, Hoxby, & Vandenbussche, 2009; Valero & Van Reenen, 2016). These benefits have encouraged individual and national investment in higher education. In fact, during the past decades, many other countries (except for China) also expanded their systems of higher education, including U.S. (see Carneiro & Lee, 2011), UK (see Murphy, Scott-Clayton, & Wyness, 2018), Russia (see Kyui, 2016), Italy (see Bianchi, 2015), Venezuela (see Gonzalez & Oyelere, 2011), and Mongolia (see Yano, 2012) among others.

These education expansion programs generally are very effective at raising students' educational attainment. However, the longer-run effects of such programs, especially their impacts on affected individuals' labor market outcomes, and more broadly, on the wage structure in the labor market at large, are much less straightforward.

In this chapter, I evaluate one large-scale and radical higher education expansion program in China to explore its impacts on students' education and labor market outcomes. The expansion was initiated in 1999 and was abrupt and unexpected for most high school students at that time.

In only three years, college enrollment nationwide doubled (MoE, 2020b). This substantially expanded access to HEIs over a short period of time provides plausibly exogenous variation in the probability of college attendance for different cohorts of students who were scheduled to attend college before or after the expansion. In other words, the expansion provides variation in college admission that is arguably unrelated to pre-existing factors like students' ability, motivation, and family background. Moreover, the intensity of the expansion was not uniform across provinces, and there were substantial inter-provincial differences in the number of increased college spots during the expansion, which provides another dimension of exogenous variation in the probability of college attendance.

Leveraging this plausibly exogenous variation in the probability of college attendance and completion, I can thus overcome non-random self-selection into higher education and therefore identify the causal effects of attending and completing college on individuals' labor market outcomes, such as the probability of being employed, work stability, and earnings. More specifically, I link province-year level administrative data on the college admission quota to a nationally-representative sample of over 10,000 adults who were born between 1959-1990 and had detailed information on their education and labor market outcomes in 2013 (when they are 23-54 years old) from the most recent round of the Chinese Household Income Project (CHIP 2013). These cohorts straddle the period during which the college expansion program was first implemented and then accelerated, and thus were differentially exposed to the expansion-induced changes in the probability of college attendance.

I then compare education and labor market outcomes for individuals who were differentially exposed to the college expansion depending on their year and province of birth. Results show that increase in educational attainment between unexposed older cohorts and exposed younger cohorts was faster in provinces that experienced greater degree of expansion. More specifically, my results show that a 10-percentage-points increase in the planned college admission ratio (equivalent to one additional college admission quota out of ten 18-year-olds) from the pre-expansion baseline level leads to 0.3 more completed years of schooling, 6-7 and 5-6 percentage points higher in the probabilities of attending college and eventually obtaining any post-secondary degrees respectively.

In addition, the plausibly exogenous increases in college admission quota also lead to increased employment probability and earnings. A 10-percentage-points increase in the planned college admission ratio improves the probabilities of working and earning positive income and working full-time-year-round by 4 and 3 percentage points respectively. When earnings are measured as hourly wage in the natural logarithm form, a 10-percentage-points increase in the planned college admission ratio links with roughly a 5-6 percent increase in hourly wages.

Combing the education and earnings effects of the college expansion suggests 2SLS estimates of the rate of return to one additional year of schooling to be approximately 18 percent, and returns to attending and completing college to be around 78-79 and 79-97 percent respectively.

I conduct a series of robustness checks to see whether the estimated education and labor market effects are sensitive to controlling for time-varying province-year level covariates, and the exclusion of two largest metropolitan cities in China. Results show that the expansion's education effects become much larger when excluding individuals from Beijing and Shanghai, and become even more pronounced after further controlling for potential post-expansion confounding influences at the province-year level. The magnitudes of the expansion's earnings effects shrink, and the coefficients generally become statistically insignificant after excluding Beijing and Shanghai and controlling for province-year level covariates.

The reminder of this paper is organized as follows. Section 2 introduces China's higher education expansion program and how I use this reform to construct my instrument for college attendance. In Section 3, I outline the data used and my empirical strategy for identifying the causal effects of reform-induced college attendance and completion on subsequent labor market outcomes. Section 4 presents results on both the quasi-experimental estimates of the expansion's effects on individuals' education and labor market outcomes, and the 2-stage least squares (2SLS) estimators of the returns to college in China. In Section 5, I conduct a series of falsification tests and robustness checks to test the validity of my identification strategy and results. Section 6 discusses major limitations of this analysis and points to the direction for future research. Section 7 concludes.

2.2 The 1999 Higher Education Expansion in China

China now has the largest higher education system in the world in terms of enrollment and degrees conferred. In the fall of 2017, China enrolled approximately 27.5 million undergraduate students in its degree-granting postsecondary institutions (National Bureau of Statistics of China [NBS], 2020), which is approximately 60 percent more than the corresponding number (16.8 million) in the U.S. (National Center for Education Statistics, 2019); By the summer of 2018, 3.8 million students earned an associate or bachelor's degree in China (NBS, 2020), roughly 30 percent more than the corresponding degrees (3.0 million) granted in the U.S. (National Student Clearinghouse, 2019).[15]

[15] Undergraduate degrees obtained for the U.S. data include those earned by both first-time graduates and graduates with prior undergraduate credentials. The China's data do not distinguish between the two.

This is a recent phenomenon, however. Just two decades ago in 1998, the number of total enrollment and degrees granted are approximately 3.4 million and 0.4 million respectively (NBS, 2020). Tertiary gross enrollment ratio is 9.8 percent in 1998 (ESYC, 1999), which is only about half of that year's world average (18.4 percent) (The World Bank, 2020a). The dramatic increases in postsecondary attendance and degree attainment can largely be attributed to China's higher education expansion that started in 1999.

In early 1999, China's central government and the Ministry of Education decided to admit approximately 470,000 more students into China's higher education institutions (HEIs) in that Fall (NBS, 2020). As a result, the total number of first-time-first-year undergraduate students newly admitted into China's degree-granting HEIs increased by nearly 50 percent between 1998 and 1999 (NBS, 2020). The radical expansion continued in the following years: number of newly admitted undergraduate students increased by about 40 percent in 2000, and kept increasing at an annual rate of around 20 percent till 2004 (NBS, 2020). In only three years from 1998 to 2001, undergraduate enrollment in China's regular HEIs more than doubled; it nearly quadrupled by 2004 (see Section 5.2 of Chapter 1 for a more detailed introduction of the policy background of the expansion program). Figure 2.1 below depicts the trends in first-time-first-year and total undergraduate enrollment in, and graduates from China's degree-granting HEIs during the years of 1995 and 2018.

This abrupt change of policy provides plausibly exogenous variation in college admission among students who were scheduled to attend college in different years. Moreover, the "intensity" in how much college admission quota is expanded is different across provinces. Under China's centralized higher education admission system, admission quotas are assigned to provinces during early months of each year, and students sit for the College Entrance Exam

(CEE) in his/her province of household registration ("hukou") during the summer after the final year of upper secondary school, and rank against each other exclusively on CEE scores within his/her cohort in his/her "hukou" province.[16]

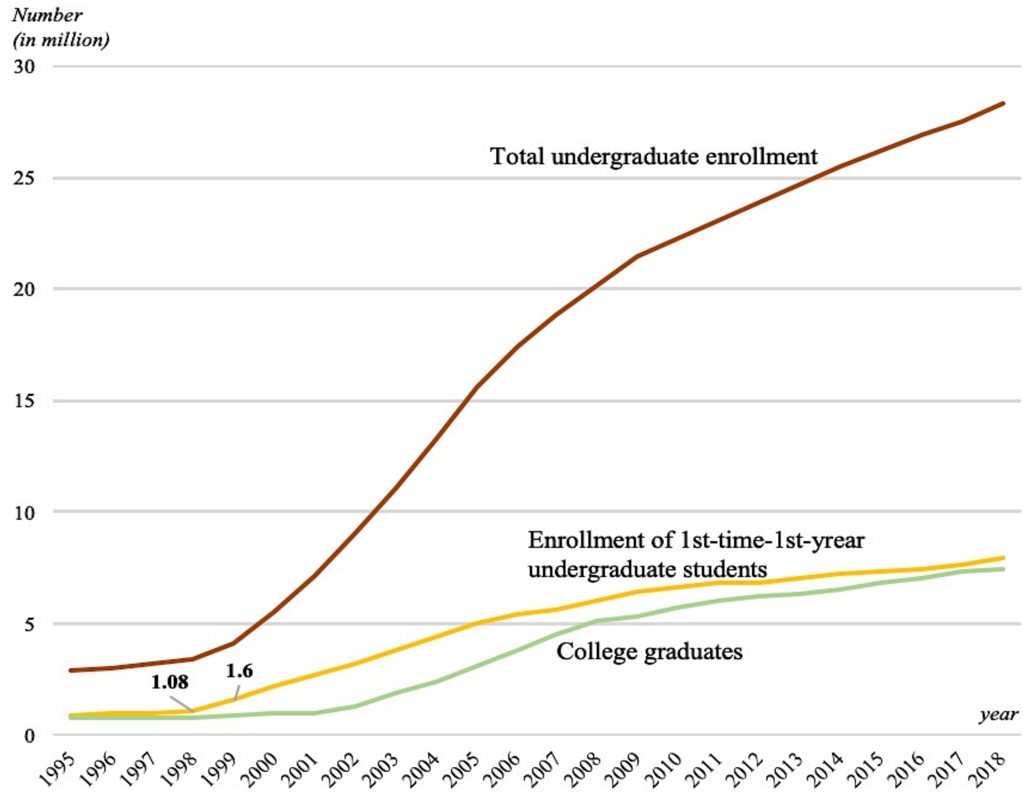


Figure 2.1: Undergraduate Enrollment in and Graduates from Regular Degree-granting Higher Education Institutions in China (in millions), 1995-2018.

Data source: National Bureau of Statistics of China (retrived from: http://www.stats.gov.cn/english/).

As a result, an individual's birth cohort and "hukou" province would together determine his or her degree of exposure to the expansion. For example, the program effects should be close

[16] "Hukou" is the household registration system in mainland China. It first classifies individuals into either an agricultural (i.e., rural) or non-agricultural (i.e., urban) "hukou." It also documents individuals' origin of birth, usually the province where one was born or the "hukou" province of one's parents. The type and province of "hukou" are linked with multiple benefits and restrictions, including where one can attend schools and take the CEE, the minimum wages and pension plans one can enjoy, and where one can purchase house and cars, etc. In most cases, the "hukou" province is equivalent to one's birth province, but children of parents who have not obtained the "hukou" of where they work/reside may have a "hukou" province equals to their parents' (birth) province.

to zero for older individuals who attended college before 1999, and largest for younger cohorts who attended college after the expansion and in provinces and years that experienced the largest increases in college admission quota.

Previous studies have used similar strategies to take advantage of the quasi-experimental nature of China's higher education expansion. More specifically, many studies (Li et al., 2014; Wu & Zhao, 2010; Xing et al., 2018; Yu, 2014) use whether one is scheduled to attend college after the expansion as an IV for college education to estimate the expansion's causal effects on employment and earnings, and generally find that the expansion lowered the average earnings of college graduates and their employment probability. One problem with this approach is that using only one dimension of variation in the probability of college attendance (i.e., due to birth cohorts) could capture the potential macro-economic trends that happened concurrently during the expansion, and misleadingly attributes them to the effects of the expansion.

Other studies try to add another dimension of variation, for example, from yearly changes in the college admission rate - defined as the number of students admitted into any HEIs divided by the number of students who took the CEE –either at the national (Hu, 2018) or province level (Ou & Zhao, 2016). These methods are better, but still have some caveats: first, under China's "quota by province" system, the national level admission rate might not be a good proxy for individuals' probability of college attendance; second, the "admission rate" measure is not exogenous because the numerator is the actual enrollment that happened after the CEE, which could be determined by students' performance on the CEE and thus cause the problem of reverse causality. The denominator is also endogenous in that expansion could induce students to take or re-take the CEE exam, or even affect high school attendance behaviors as well. Later in section 3, I will show that admission quota set before the CEE and the number of 18-year-old population could be a better choice to construct the admission probability measure. For a more thorough review of prior literature on the effects of China's higher education expansion, see Section 5.3 in Chapter 1.

2.3 Data and Method

2.3.1 Data

Data used in this analysis come from two major sources: the individual level data come from the Chinese Household Income Project (CHIP), and the province-cohort level college admission quota data are constructed using administrative records from both the Department of Development and Planning under Ministry of Education, and from various years of the *Educational Statistics Yearbook of China*. Other sources of data include various years of Census data and yearbooks, obtained from the National Bureau of Statistics of China's data archive (NBS, 2020).

The CHIP is a repeated cross-sectional household survey that has been conducted in 1988, 1995, 2002, 2007 and 2013. It uses multi-stage random sampling to draw nationally representative samples of households, which are sub-samples of the National Bureau of Statistics' larger official annual household surveys (Gustafsson, Li, & Sato, 2014). The purpose of the survey is to measure and estimate the distribution of personal income in China, and thus the survey contains very detailed information on employment (including number of jobs held, month, weeks and hours worked, industry, occupation, position, and type of work, etc.), and various income measures (including all sources of incomes, and job benefits). It also contains detailed information on educational history (including years of schooling completed, highest educational attainment, type of high school and college attended, and information about CEE taking, etc.). Importantly, the dataset contains sufficient information on individuals' year and month of birth, and province of birth that can be used to clearly identify the "intensity" of expansion an individual was exposed to. Finally, the data include demographic information on gender, type of household registration at birth (i.e., rural or urban), ethnicity, and current province of residence/work.

To estimate the impacts of the college expansion on education and earnings, I use the most recent round of CHIP (i.e., CHIP 2013), which covers earnings data in 2013 and could thus capture labor market outcomes for ten post-expansion cohorts. The CHIP 2013 sample also has the broadest geographical coverage that contains individuals currently residing in 14 of the 32 provincial districts in mainland China, and their "hukou" province represent 31 of the 32 provinces (only missing individuals from Hainan province).[17] Later in Section 5, I also use the two prior-expansion waves (CHIP 1988 and CHIP 1995) to test for the parallel pre-trends hypothesis. I restrict my analysis to the sub-sample of urban residents for two major reasons: first, primary source of income for most rural residents comes from leasing land or agricultural work that could be very different from the employee-type of jobs for urban residents, and second, there are very low percentage of rural residents who ever had post-secondary education. The CHIP 2013 urban sample contains 19,887 individuals.

2.3.2 Sample

The "full analysis sample" includes individuals born between September of 1959 to August of 1990 (i.e., the 1978-2008 college cohorts). Individuals older than the 1979 college cohort were affected by the "Culture Revolution" during which all the higher education

[17] The sampling frame and questions change frequently from wave to wave for the CHIP datasets. Other reasons for using the CHIP 2013 dataset include only this wave contains CEE-related questions, including year and province taking the CEE that is crucial to determine treatment intensity. Moreover, previous waves of the CHIP datasets cover at most 16 provinces (compared with 31 provinces for the CHIP 2013 dataset), which would substantially reduce the range of variation in province-year-level treatment intensity.

institutions were shut down (older cohorts in the survey are also likely to be retired by 2013); Meanwhile, individuals younger than the 2008 college cohort are likely to be still enrolled in college or graduate schools by 2013, creating downwards bias if excluding those enrolled in school. I further exclude individuals who did not at least complete lower-secondary school by 2013 since China has the "9-year compulsory schooling" laws (corresponding to the completion of lower-secondary school) and only approximately 5 percent of the full sample did not complete lower-secondary school.[18] I further exclude those who were enrolled in school at the time of the survey, whose post-secondary degrees were obtained from non-regular higher education institutions (e.g., party schools, adult education or other part-time programs), and those who has obviously misreported their educational attainment.[19] Finally, I exclude the 81 individuals from the transition cohorts (i.e., those who would either attend college in 1998 just before the expansion or in 1999 just after the expansion) whose probable college cohorts cannot be inferred from the information they provide since it is hard to determine the "treatment" status for these individuals. The remaining analysis sample includes 8,379 individuals.

To examine the earnings effect of the college expansion, my "worker" sample restricts to those who reported to have worked for at least one month and received positive income from any type of job during 2013. The worker sample consists of 7,108 individuals, accounting for 84.8 percent of the full analysis sample.

[18] The law was implemented in 1986, and the oldest elven cohorts in this sample could drop out of lower secondary school before it. Nevertheless, the overall proportion of individuals who have not completed at least nine years of education is low.

[19] Those people usually "attend" postsecondary programs after many years of working, and obtain such certificates purely for promotion purposes. They are very likely to be different than the traditional college students who attend regular HEIs right after high school.

2.3.3 Estimation Strategy

The key assumption underlying my identification strategy is that the post-expansion variation in the provincial-year level increases in college admission quota (from their respective pre-expansion levels) are uncorrelated with potential unobserved province-year level confounding factors (such as demand for college education or demand for high-skill labor in the local labor market) that also influence students' educational and subsequent labor market outcomes. If this assumption holds, then we can use this variation to causally evaluate the effects of the expansion, by comparing whether younger cohorts have better educational and labor market outcomes relative to older cohorts in the provinces and/or years that received higher expansion-induced increases in the college admission quota. Note that I use the ***increases*** in the number of provincial-year level college admission quota from their corresponding pre-reform level, not the absolute number of quota (which could be highly related to confounding factors that might also simultaneously affect individuals' future education and labor market outcomes).

In Section 5, I formally test the validity of my identifying assumption that the province-year level increases in college admission quota after the expansion are exogenous. I further relax this assumption by exploring and controlling for the factors that might determine both these increases and potential outcomes, thus only using the part of exogenous variation in the increases in college admission quota induced by the expansion to remove any potential confounding influences. This is explained in more details in the "robustness checks" in Section 5.

The reduced-form estimating equation can be written as a two-way fixed-effects model with controls:

$$Y_{i,p,c}=c_1 + \alpha \left(\boldsymbol{Post_c} * \boldsymbol{Intensity_{p,c}}\right) + \beta \boldsymbol{Cohort}_c + \gamma \boldsymbol{Province}_p + \theta \boldsymbol{X}_i + \varepsilon_{i,p,c}, \quad (2.1)$$

where $Y_{i,p,c}$ is the education or labor market outcomes for individual *i* from province *p* and probable college cohort *c*; $\boldsymbol{Post_c}$ indicates whether individual *i* would attend college after the expansion; $\boldsymbol{Intensity_{p,c}}$ denotes province-cohort level intensity of the expansion, measured as the increase in admission quotas to regular HEIs (both 3- and 4-year colleges) assigned to province *p* for cohort *c* from the base level in 1998 for province *p*, both adjusting by the number of 18-year-olds at the province-cohort level; $\boldsymbol{Cohort_c}$ is probable college cohort, inferred from individuals' year and month of birth assuming that individuals attend college at 18 years old; $\boldsymbol{Province_p}$ is the province where individuals took the CEE, or the "hukou" province at 18 years old for those who did not take the CEE; $\boldsymbol{X_i}$ is a vector of individual characteristics, including gender, a rural/urban dummy, and an interaction between the two variables; $\varepsilon_{i,p,c}$ is the error term. I provide more details regarding definition and construction of key variables in the next section to follow.

The variable of interest is the coefficient before the interaction term between "$\boldsymbol{Post_c}$" and "$\boldsymbol{Intensity_{p,c}}$" that identifies the variation in expansion intensity among exposed cohorts. This coefficient (i.e., α) represents whether the increases/decreases in outcomes observed for exposed cohorts relative to unexposed cohorts are larger for those from provinces that experienced larger increases in college admission quota.

2.3.4 Key Variables

Dummy variable for before- and after-expansion cohorts. The CHIP 2013 dataset contains individuals' month and year of birth, which I use to infer individuals' "probable college cohort" assuming that individuals start first grade of primary school after they turn 6 and would attend college when they are between 18 to 19 years old.[20] I then construct the dummy variable, $\boldsymbol{Post_c}$, that indicates whether students are "scheduled" to attend college after the expansion. In other words, $\boldsymbol{Post_c}$ equals 1 for those with probable college cohorts of 1999 and after, and equals 0 for those of prior-1999 probable college cohorts.

The CHIP 2013 dataset does not contain information on the year of college attendance, but has information on the most recent year of taking the CEE (though with a high percentage of missing on this variable). For those who have valid data on this variable, approximately 39 percent of individuals took the CEE at the age of 18-19 and thus would attend college in the year exactly as what the "probable college cohort" variable would predict; another 42 percent of individuals took the exam at age 17-18 or 19-20, that is, they would attend college one year before or after the "probable college cohort" variable would predict. This imperfect correspondence could be due to some students skipping grades, students entering primary school early or later than expected, or that some students repeated the last year of senior high school and re-took the CEE. Given this non-negligible proportion of individuals who either take the CEE one year before or after the inferred "probable college cohort" year, I exclude the two transitory cohorts who would attend college right before or right after the expansion (i.e., whose probable cohort of college attendance is 1998 or 1999) to avoid high measurement error in determining their treatment status. The two transitory cohorts account for 3.6 percent of the entire sample.

Expansion intensity. As introduced in Section 2, students are only eligible to take the CEE at their "hukou" province, and thus the expansion intensity of that province at his/her

[20] More specifically, in most provinces in China, it is required that students can start primary school (usually from September 1st) when they reach 6 on August 31st in that year. It takes 12 years to complete upper-secondary school before one can enter college (in the majority of provinces in China, primary school takes 6 years, and lower- and upper-secondary schools take 3 years respectively). For example, individuals with a probable college cohort of 2008 born between September 1, 1989 and August 31, 1990.

probable college cohort year would determine the degree the individual was exposed to the expansion. The CHIP datasets ask about province of current residence, province of current "hukou," and province where one took the CEE. For the last one, only 31.6 percent of the full sample from the CHIP 2013 dataset reported the province where they took the CEE. For those who have missing values on this variable, I use their "hukou" province at the time of survey to infer their "hukou" province when they were about to take the CEE. Individual could change their "hukou" province after 18 years old, so the "hukou" province in 2013 might not accurately reflect the province where an individual would take the CEE. Though I cannot entirely dismiss this possibility, I show that mobility across provinces or change of "hukou" are relatively rare for the CHIP 2013 dataset: among those who have valid data on the province where they took the CEE, only 2.8 percent out of the full sample has a CEE province different from their "hukou" province in 2013, and only 2.4 percent reside or work in a province different from their "hukou" province in 2013. Nevertheless, I later address the problem of selective migration either due to migrating to another province to take the CEE or to work later in Section 5.

I then determine an individual's degree of exposure to the expansion by constructing an "intensity" measure. I first construct the province-cohort level college admission ratio as $Ratio_{p,c} = \frac{Quota_{p,c}}{18-year-olds_{p,c}}$. The numerator is the number of planned admission quota assigned to students from CEE province p in probable college cohort c. I use the quota, which is set before students taking the CEE, rather than the actual enrollment, which is later determined by both students' test scores and their application choices.[21] The denominator is the number of 18-year-old population in province p and cohort c. I then use the increases in $Ratio_{p,c}$ for the post-

[21] Table A1 in Appendix A indicates, the number of college admission quota and actual new college enrollment are roughly the same at the national aggregate level for most years during 1995-2010.

expansion years from the corresponding province level ratios in 1998 to determine expansion intensity, that is, $Intensity_{p,c} = Ratio_{p,c} - Ratio_{p,1998}$ for c >1998, and $Intensity_{p,c} = 0$ for all the before-expansion cohorts.

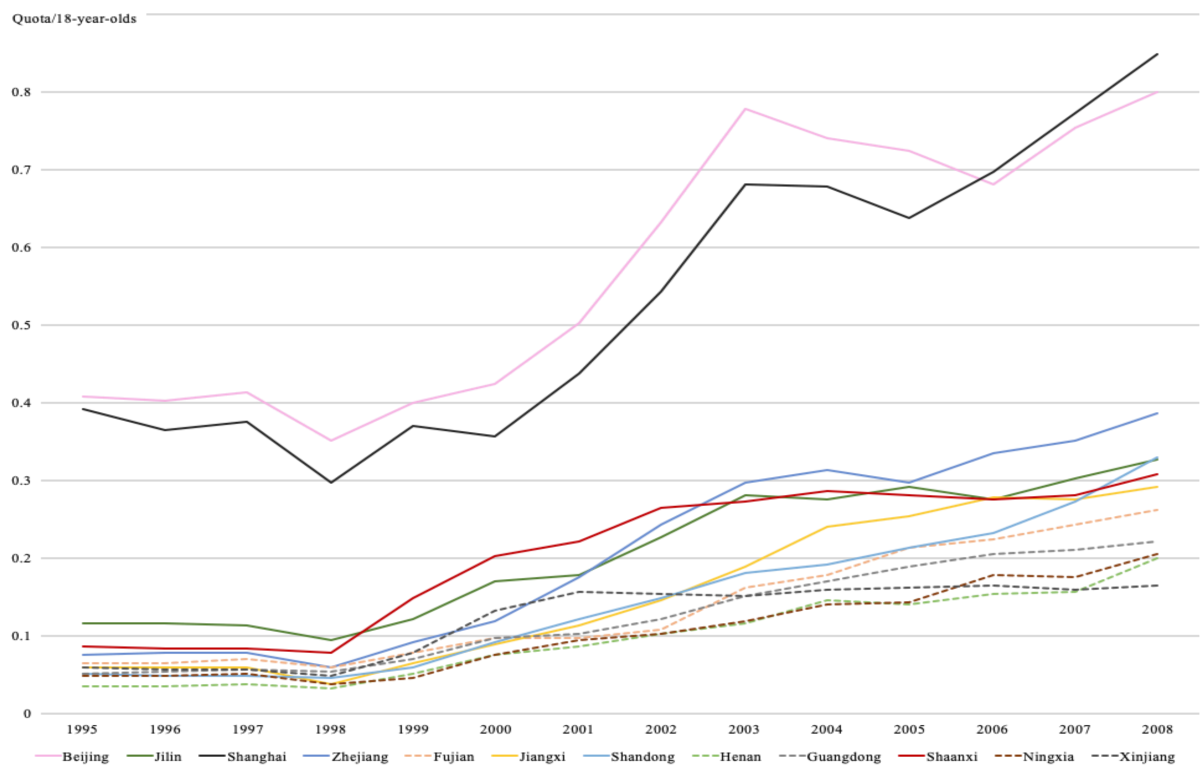
For the denominator, I compute province-cohort level college-age population using the Census data from NBS (2020). The Census data have information on the national aggregate population at each age for each year, but at the province level, they only contain population in the 15-19 age group. I infer the size of 18-year-olds using its proportion out of the 15-19 age group in the national aggregate sample, assuming that the distribution of each age cohort within the 15-19 age group are the same across all provinces.[22] Alternative candidates for the denominator such as the number of CEE-takers or high school graduates at the province-cohort level are more problematic because test-taking and high-school attendance are endogenous and likely to be affected by the expansion.

The CHIP 2013 sample covers 30 CEE provinces (out of the total of 31 provinces in mainland China). I construct a panel of provincial college admission ratio for all these 30 provinces from 1995 through 2008 mainly using administrative record from the Ministry of Education's Department of Development and Planning (presented in Table A2 in Appendix A). The last column under each post-expansion years indicate whether the province-year level expansion intensity is above ("H") or below ("L") the nationwide "intensity" for that year.

Panel A of Figure 2.2 below presents trends in college admission ratios for half of the total 30 provinces (covering all the five larger geographic regions in China). Panel B of Figure 2.2 presents the same results excluding the two outlier cities - Beijing and Shanghai.

[22] I also use two alternative imputation strategies (i.e, using province-year level number of primary school and lower secondary school graduates six and three years before the "probable college cohort year" multiplied by their corresponding gross enrollment rates at respective educational levels), and derive similar numbers.

A. Including Beijing and Shanghai



B. Excluding Beijing and Shanghai

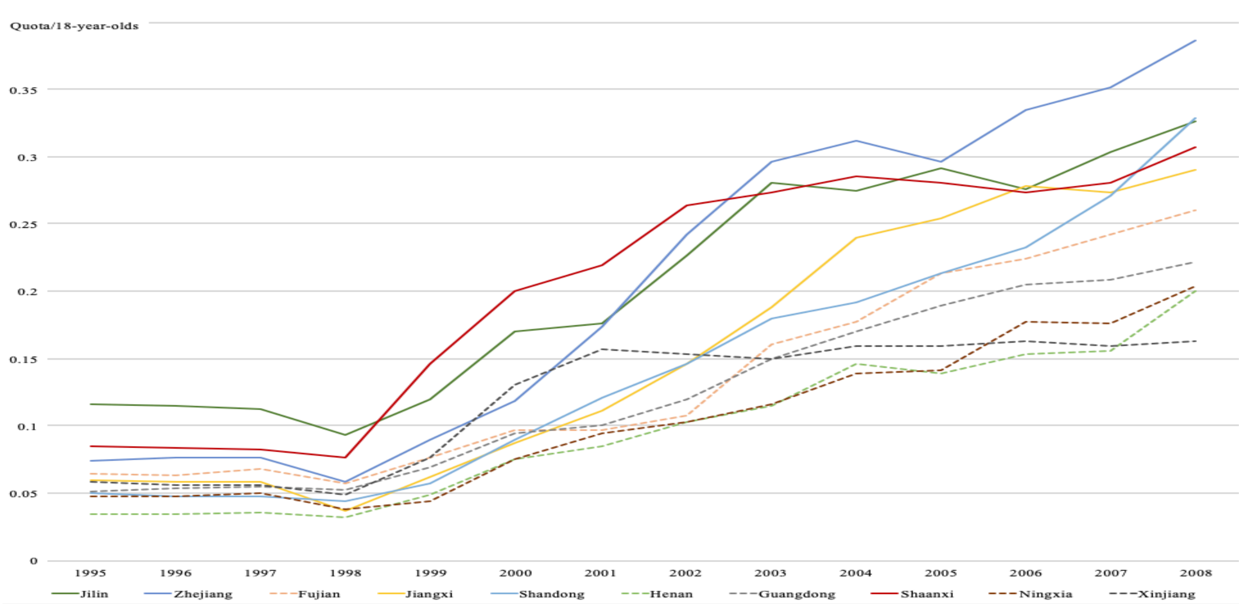


Figure 2.2: College Admission Ratio for Selected Provinces, Including and Excluding Beijing and Shanghai, 1995-2008.

Data source: Author's calculation using administrative data from MoE's Department of Development and Planning.

Figure 2.2 suggests that trends in college admission ratio at both the nationwide and provincial level were quite stable during 1995 to 1997 except for a relatively uniform dip in 1998, and the trends are parallel across provinces. Provinces experienced more substantial changes in admission ratio starting from 1999 when the higher education expansion was first initiated, and province-level trends after the expansion became more idiosyncratic. Moreover, the intensity measures for the panel of 300 observations (30 provinces * 10 post-expansion cohorts) range from 0.011 to 0.490, with a mean of 0.140 for the full sample and a standard deviation of 0.09. These statistics indicate that during all the years after the expansion, all the provinces increase college admission spots to some extent, ranging from increases of 11 to 490 additional spots out of a thousand 18-year-olds.

Comparing across provinces, the means of this "intensity" measure for all the province-cohort level high- and low-intensity observations are 0.201 and 0.102 respectively, and the difference between the means is roughly 10 percentage points. This implies that, on average, one additional person out of ten 18-year-olds would be admitted into any regular HEIs in China during the initial ten years after the expansion in high-intensity provinces as compared with people in low-intensity provinces.

Moreover, when comparing across years, at the national level, the "intensity" measure for the year of 2008 (i.e., the youngest college cohort in my sample) is 0.202 from the baseline level of nationwide college admission ratio of 0.056. This indicates that in 1998, for every 1,000 18-year olds, on average, 56 people would be admitted into any regular higher education institutions in China; In 2008, 202 more people out of the one 1,000 18-year-olds (that is roughly one extra person out of every five 18-year-olds) would be admitted into college. Lastly, Figure 2.2 indicates that Beijing and Shanghai have far higher ratios in college admission both before and after the expansion than other provinces, and due to this reason, I later perform all the analyses excluding individuals from these two metropolitan cities as robustness checks in section 5.

Outcomes. I construct three variables to measure the educational effects of the expansion: the "years of schooling completed" variable measures the highest educational level one completed by the end of 2013, and six, nine and twelve years of schooling corresponds to the completion of primary, lower- and upper-secondary schools; The other two dummy variables are used to capture college attainment: the "ever attended regular post-secondary institutions" dummy equals one if individuals ever attended any regular HEIs, including 3-year community colleges, 4-year universities, and graduate schools; the "obtained any post-secondary degrees" dummy equals one if individuals ever obtained an associate's (AA), bachelor's (BA) or graduate degrees from regular HEIs. I only focus on degrees obtained from regular HEIs. Those who attended adult HEIs might obtain certificates or diplomas that in principal are equivalent credentials, although the standards for these degrees and the signals they carry fall short of the standards for AA, BA or graduate degrees from full-time regular HEIs.

As for labor market outcomes, I construct outcomes for employment probability and stability, quantity of labor supply, and earnings. The "probability of working and earning positive income" variable equals one if an individual worked for at least one month and received positive income from any type of job during 2013, and the "probability of working for a regular type of job" variable equals one if an individual worked as an employee and thus excludes the following types of employment: self-employment, work in home, agricultural work (mostly leasing land), and work as employers. The "worked full-time and year-round" dummy variable equals one if an individual worked for 12 months and on average 35 or more hours per week during 2013;[23] the “held a long- or fixed-term contract” dummy variable equals one if an individual worked under a long- or fixed-term employment contract, and equals zero if worked under a short-term (less than 1 year), contemporary, or non-formal employment contract; the “number of jobs held” variable measures the total number of jobs held during 2013; the “total hours worked for all jobs” variable is calculated as the sum of self-reported total weeks worked multiplied by total hours worked per week for all the jobs held in 2013. For individuals in the full analysis sample who reported to have not worked during 2013, I assign zero values to these employment probability and stability outcomes, and thus include them in the analysis.

With regards to income, I construct four measures to capture both the total annual income and hourly wage in 2013 expressed in both the level and natural logarithm terms. More specifically, the “total annual income from all jobs” variable capture the sum of incomes from all jobs held in 2013, and incomes include salary, bonus, and in-kind job benefits such as discounted housing and food vouchers (transferred to Chinese yuan values). The “hourly wage from all jobs” variable is the corresponding total annual income divided by the total hours worked in 2013, and “ln(total annual income)” and “ln(hourly wage)” are the corresponding income measures in the natural logarithm form. For individuals who never worked during 2013, I assign them zero values to all the income variables measured in the natural logarithm terms, and use imputed log earnings by adding 1 RMB to hourly wage for all individuals in the full analysis sample. I also add a dummy variable indicating whether an individual has zero earnings later in all the earnings regressions when using imputed earnings.

[23] National holidays, annual leave, vacations for certain jobs (e.g., summer and winter vacations for teachers) are regarded as working days in the CHIP 2013 survey.

Covariates. Individual level covariates include a "female" dummy, a "rural" dummy indicating the type of an individual's "hukou" (it equals zero when an individual holds an urban "hukou"), and an interaction between the two variables.

Table 2.1 below provides summary statistics for the full and worker samples. There are 8,379 individuals in the full sample, of whom approximately 27.2 percent belong to the post-expansion cohorts. On average, they have 11.6 years of completed schooling (12 years of schooling correspond to the completion of high school). Approximately 32 percent of individuals in the full sample have attended postsecondary education, and 29 percent have obtained a postsecondary degree. In 2013, the average age of individuals in the full sample is 39 years old, and 84.8 and 72.8 percent ever worked and worked for an employee-type job in 2013 respectively.

The worker sample is slightly more selective in terms of educational attainment: the means for years of completed schooling are 11.8 years, and higher proportions of individuals have attended postsecondary education and obtained any postsecondary degree as well. For those who ever worked and earned a positive income in 2013, most individuals (98.6 percent) worked for only one job in 2013; roughly 77 percent worked full-time and year-round in 2013, and 56 percent were employed under a long- or fixed-term contract. In terms of earnings, the mean of total annual income from all jobs in 2013 are 41,076 Chinese yuan (CNY) for the worker sample; When taking zero earnings for those who did not work in 2013 into consideration, the mean of total annual income from all jobs in 2013 are 34,643 CNY for the full sample. These are approximately \$6,455 and \$5,444 in 2018 US dollars respectively (inflated to 2018 CNY first and converted to USD using an exchange rate of 1 USD=7 CNY). The hourly wage for the worker and full samples are 21 and 18 CNY, corresponding to \$3.3 and \$2.8 in 2018 US dollar.

Table 2.1: Descriptive Statistics for the Full and Worker Samples.

Variable	Full Analysis Sample (N=8,379) Mean	Workers Sample (N=7,108) Mean
Education Outcomes in 2013:		
Years of schooling completed	11.6	11.8
Ever attended regular post-seconary institutions	0.320	0.351
Obtained any postsecondary degrees	0.291	0.322
Completed lower-secondary school only	0.335	0.306
Attended high school without obtaining any HS diploma	0.004	0.004
Obtained an vocational high school diploma only	0.139	0.145
Obtained an academic high school diploma only	0.231	0.223
Attended college without obtaining any post-seconary degrees	0.029	0.031
Obtained an associate's degree only	0.120	0.128
Obtained a bachelor's degree only	0.150	0.169
Obtained any graduate degrees	0.022	0.025
Labor Market Outcomes in 2013:		
Ever worked and earned a positive wage	0.848	---
Earned positive wages from employee-type jobs	0.728	0.858
worked full-time and year-round	0.653	0.769
held a fixed- or long-term contract	0.465	0.561
Number of jobs held	0.86	1.02
Total annual income from all jobs (¥)	34,643	41,076
Total hourly annual income from all jobs (¥)	18.0	21.0
Log (total annual income)	---	10.36
Log (total annual income), imputed	8.82	---
Log (hourly total annual income)	---	2.66
Log (hourly total annual income), imputed	2.32	---
Instrument:		
"After-expansion" cohorts	0.272	0.278
Expansion intensity for "after-expansion" cohorts	0.142	0.142
Control Variables:		
Female	0.506	0.452
Age	39.3	38.9
Rural "Hukou"	0.400	0.399

Data source: Chinese Household Income Project 2013.

2.4 Results

I first present results from the baseline models that do not include the interactions between the "$\boldsymbol{Post_c}$" dummy and province-year level covariates (i.e., $\boldsymbol{Post_c * Covariates_p}$), and later present results with these additional controls as robustness checks in section 5.

Intuition behind my estimating equation (2.1) can be illustrated using a simple two-by-two table. I divide the 300 post-expansion province-year level observations of expansion intensity into either the "high" or "low" groups, and Table A3 in the Appendix shows the means of education and labor market outcomes for the prior- and post-expansion cohorts and the high- and low-intensity groups for the worker and full samples respectively.

In both low- and high-intensity groups, younger exposed cohort have better educational outcomes (compared with older unexposed cohorts), and the improvements are larger for individuals in the high-intensity group. As for earnings, in both low- and high-intensity groups, younger exposed cohort has lower hourly wage (compared with older unexposed cohorts), but the gap between the old and young cohorts is smaller for individuals in the high-intensity group. Moreover, for the full sample, younger post-expansion cohorts in the high intensity group are also more likely to be working and earning positive earnings.

The differences in differences (DiD) presented at the last columns in Table A3 can be interpreted as the causal effects of the expansion on education and earnings, under the assumption that without the expansion, changes in educational attainment and earnings during the before- to post-expansion period would not be systematically different in the high- and low-intensity groups. Using the worker sample as an example, the magnitudes of the DiD estimators imply that for individuals young enough to attend college after the expansion and belongs to provinces and cohort years that experienced above national-average increases in college admission quotas, they received on average 0.4 more years of schooling and have 7.5 and 6.0 percentage points higher probabilities of attending college and obtaining any postsecondary degrees respectively. Their hourly wage is also 3.4 percent higher.

Instead of using a dichotomous measure of the $\boldsymbol{Intensity_{p,c}}$ (i.e., either "high" or "low"), I can take advantage of the full variation in the province-cohort level expansion intensity. Below I present regression results of the expansion's causal effects on education and labor market outcomes using my estimating equation (2.1) with a continuous measure of $\boldsymbol{Intensity_{p,c}}$ that exploits variation in treatment intensity across provinces and cohort years.

2.4.1 The College Expansion's Effects on Education

If the interaction between the "post" and "intensity" measure validly identifies exogenous variation in college education, students from provinces that experienced larger increases in the college admission quota due to the expansion should have better educational outcomes compared with those from provinces that had lower expansion intensity. Table 2.2 presents regression results of the effects of the college expansion on education outcomes using equation (2.1) for the full and worker sample respectively.

Table 2.2: Estimating the Effects of the Expansion on Education Outcomes.

Educational Outcomes	Full Sample *(N=8,377)*				Workers Sample *(N=7,106)*			
	[1]		[2]		[3]		[4]	
	raw coefficient		*10 pp. increase in "intensity"*		raw coefficient		*10 pp. increase in "intensity"*	
Years of schooling completed	3.4	***	*0.3*	***	2.7	**	*0.3*	**
	(0.9)		*(0.1)*		(1.0)		*(0.1)*	
Attended any post-seconary institutions	0.715	***	*0.072*	***	0.596	***	*0.060*	***
	(0.152)		*(0.015)*		(0.167)		*(0.017)*	
Obtained any postsecondary degrees	0.606	***	*0.061*	***	0.468	**	*0.047*	**
	(0.148)		*(0.015)*		(0.164)		*(0.016)*	

Data source: Chinese Household Income Project 2013.

Notes: Robust standard errors clustered at the province level are in parentheses.

***, **, and * indicate significance at the 0.1, 1, and 5 percent levels.

Using variation in the expansion-induced increases of admission quota across cohorts and provinces, results show that the expansion boosts individuals' educational outcomes, including years of completed schooling, and the probabilities of attending and completing college. Note that at the national level, the admission ratio increased by approximately 20 percentage points by 2008 from the baseline level of 5.6 percent in 1998 (i.e., before the expansion started). The difference in means of the high- and low-intensity groups is approximately 10 percentage points. Therefore, besides presenting the raw coefficients on the interaction term between intensity and the post dummy in columns 1 and 3 under the two samples respectively, in columns 2 and 4 I also present the effects of a 10-percentage-point increase in expansion intensity (equivalent of 1 additional college admission spot out of 10 18 year-olds after the expansion), which is also roughly the difference in the means of expansion intensity between the high- and low-intensity groups. For example, for the full analysis sample, a 10-percentage-point increase in expansion intensity led to approximately 0.3 more years of completed schooling, 7.2 percentage points higher probability of attending college, and 6.1 percentage higher probability of obtaining any post-secondary degrees. In other words, the growth in completed years of schooling from unexposed older cohorts to exposed younger cohorts is 0.3 years higher for those from the high-intensity group (as compared with those from the low-intensity group); the growth in the probabilities of attending any post-secondary institutions and obtaining any postsecondary degrees are also 7.2 and 6.1 percent higher for those from provinces and years that experienced higher expansion intensity than the corresponding nationwide average. All coefficients on the educational outcomes are statistically significant at the 0.001 level.

The third and fourth columns present the same sets of results for those who reported positive income in 2013. The pattern of result is very similar to those for the full analysis sample, however, the magnitudes of the coefficients for the three educational outcomes are smaller since individuals in the worker sample have higher educational attainment to begin with. Estimates of the expansion's educational are statistically significant at least at the 1 percent level for the worker sample.

I also construct more detailed categories of educational levels to explore at which level the expansion's educational effects are the most pronounced. The expansion's effects on education could work through the following channels: first, it could increase high school graduates' chances of being admitted into college, and thus increase their probability of college attendance and attainment; and second, it could also induce students to finish high school or attend high school (even if they do not eventually end up attending any post-secondary institutions), and thus increase their educational attainment at the secondary school level. Results (presented in Table A4 in Appendix A) suggest that the expansion's effects on education mainly work through inducing academic high school graduates who would directly enter the labor market without attending any college to attend and complete college and even beyond.

2.4.2 The College Expansion's Effects on Labor Market Outcomes

I then tend to the expansion's effects on individuals' labor market outcomes in 2013. Table 2.3 presents regression results from equation (2.1) for both the full and worker samples. The outcomes are defined as in section 3.

For the full sample, Table 2.3 shows that the expansion has positive impacts on the probabilities of working and earnings positive income from both any type and employee-type of jobs. More specifically, a 10-percentage-point increase in the intensity of the expansion is associated with approximately 4-percentage-point increases in the probabilities of ever worked for positive income in 2013 and working for an employee-type job. The coefficients are statistically significant at least at the 0.01 level. The expansion also improves individuals' probability of working full-time and year-round: a 10-percentage-point increase in the intensity of the expansion is associated with approximately a 2-percentage-point increase in the probabilities of working full-time and year-round during 2013. When using imputed income that also take into consideration of zero earnings for the non-workers, the expansion also has a positive impact on hourly wage: a 10-percentage-point increase in the college admission quota boosts hourly wage in 2013 by 5.7 percent. This coefficient is significant at the 5 percent level.

Table 2.3: Estimating the Effects of the Expansion on Labor Market Outcomes.

Labor Market Outcomes	Full Sample *(N=8,377)* [1] raw coefficient		[2] *10 pp. increase in "intensity"*		Workers Sample *(N=7,106)* [3] raw coefficient		[4] *10 pp. increase in "intensity"*	
Probability of working and earning positive income	0.376	***	*0.038*	***	---		---	
	(0.117)		*(0.012)*		---		---	
Probability of working for an employee-type of job	0.414	**	*0.041*	**	0.026		*0.003*	
	(0.150)		*(0.015)*		(0.129)		*(0.013)*	
Probability of working full-time-year-round	0.343	*	*0.034*	*	0.010		*0.001*	
	(0.160)		*(0.016)*		(0.155)		*(0.016)*	
Probability of working under a long-term employment contract	0.172		*0.017*		-0.072		*-0.007*	
	(0.172)		*(0.017)*		(0.187)		*(0.019)*	
(imputed) log(hourly total income)*	0.574	*	*0.057*	*	0.532	~	*0.053*	~
	(0.226)		*(0.023)*		(0.294)		*(0.029)*	

Data source: Chinese Household Income Project 2013.

Notes: The full sample uses imputed log hourly income by adding 1 RMB to all individuals and adding a dummy variable that equals one for non-workers. Robust standard errors clustered at the province level are in parentheses.

***, **, *, and ~ indicate significance at the 0.1, 1, 5, and 10 percent levels.

For individuals who have earned positive earnings during 2013 (i.e., the "worker sample"), the college expansion increases their hourly wage: a 10-percentage-point increase in the college admission quota boosts hourly wage in 2013 by 5.3 percent. This coefficient is marginally significant at the 10 percent level. The expansion does not have statistically significant impacts on the probabilities of holding an employee-type job, working full-time and year-round, or working under a long-term employment contract for the worker sample.

2.4.3 Estimating the Causal Returns to College

Finally, I estimate the effects of expansion-induced college attendance and completion on employment and earnings. After controlling for cohort and province fixed effects, the interaction between dummy variables indicating whether individuals attend college after the expansion (i.e., $\boldsymbol{Post_c}$) and the province-cohort level expansion intensity measure (i.e., $\boldsymbol{Intensity_{p,c}}$) are a plausible exogenous variable for education. In practice, I use the interaction (i.e., $\boldsymbol{Post_c} * \boldsymbol{Intensity_{p,c}}$) as an instrument for the following three endogenous variables: years of schooling completed, the probability of attending any post-secondary institutions, and the probability of obtaining any post-secondary degrees, and use the following two-stage least squares (2SLS) models to estimate the causal returns to an additional year of schooling, college attendance, and post-secondary degree attainment on labor market outcomes.

$$\widehat{Edu}_{i,p,c} = c_1 + \pi_1(\boldsymbol{Post_c} * \boldsymbol{Intensity_{p,c}}) + \pi_2 \boldsymbol{Cohort}_c + \pi_3 \boldsymbol{Province}_p + \pi_4 \boldsymbol{X}_i, \quad (2.2)$$

$$Y_{i,p,c} = c_2 + \alpha * \widehat{Edu}_{i,p,c} + \beta \boldsymbol{Cohort}_c + \gamma \boldsymbol{Province}_p + \theta \boldsymbol{X}_i + \varepsilon_{i,p,c}. \quad (2.3)$$

Dependent variable in equation (2.2) ($\widehat{Edu}_{i,p,c}$) is predicted education outcome due to the treatment (i.e., expansion-induced increases in college admission spots), including years of completed schooling, and the probabilities of college attendance and post-secondary degree attainment. Dependent variable in equation (2.3) ($Y_{i,p,c}$) is labor market outcome in 2013, including the probabilities of working and earning positive income, working for an employee-type of job, and working full-time-year-round, and log annual total income. All other variables are defined in the same way as in equation (2.1). Note that the 2SLS estimator (α) of the effects of education on labor market outcomes is roughly the ratio between the reduced-form estimator of the expansion's labor market effect divided by the first-stage estimator of the expansion's effect on education.

Table 2.4 below presents results from the 2SLS regressions that use post-expansion intensity as the instrument for the three educational outcomes. More specifically, Models 1 and 2 present 2SLS estimates of the effects of one additional year of schooling on labor market outcomes, such as employment probabilities, employment stability, and earnings, for the full and worker samples respectively. Models 3 and 4 present 2SLS estimates of the effects of college attendance, and models 5 and 6 present 2SLS estimates of the effects of postsecondary degree attainment, all on the same sets of outcomes.

Table 2.4: 2SLS Estimates of Expansion-induced Increases in Education on Labor Market Outcomes.

Outcome	[1] Full Sample *(N=8,377)*	[2] Worker Sample *N=7,106*	[3] Full Sample *N=8,377*	[4] Worker Sample *N=7,106*	[5] Full Sample *N=8,376*	[6] Worker Sample *N=7,105*
	Instrumented: Years of schooling completed		*Instrumented: Attended any postseconary institution*		*Instrumented: Obtained any postsecondary degree*	
Probability of working and earning positive income	0.112 **	--- ---	0.526 **	--- ---	0.621 **	--- ---
	(0.043)	--- ---	(0.187)	--- ---	(0.231)	--- ---
Probability of working for an employee-type of job	0.124 *	0.010	0.580 **	0.044	0.700 **	0.079
	(0.050)	(0.046)	(0.222)	(0.213)	(0.272)	(0.271)
Probability of working full-time-year-round	0.102 *	0.004	0.480 *	0.017	0.564 *	0.018
	(0.051)	(0.056)	(0.232)	(0.259)	(0.278)	(0.330)
(imputed) log(hourly total income)*	0.179 *	0.177 ~	0.783 **	0.788 ~	0.914 *	0.968 ~
	(0.073)	(0.096)	(0.304)	(0.421)	(0.364)	(0.532)

Data source: Chinese Household Income Project 2013.

Notes: Robust standard errors clustered at the province level are in parentheses.

***, **, *, and ~ indicate significance at the 0.1, 1, 5, and 10 percent levels.

Models 1, 3, and 5 of Table 2.4 show that for the full sample, improvement in education also boost employment outcomes: an additional year of schooling increases the probabilities of working and earning positive income, working for an employee-type of job (compared with self-employment), and working full-time-year-round by 11.2, 12.4, and 10.2 percent respectively. Attending and completing college has larger positive impacts on employment: effects of attending college and obtaining any post-secondary degree on the probability of working and earning positive income are 52.6 and 62.1 percent respectively; their effects on the probability of working for an employee-type of job are 58.0 and 70.0 percent respectively, and they also increase the probability of working full-time and year-round by 48.0 and 56.4 percent respectively. Coefficients on these employment outcomes are statistically significant at least at the 5 percent level. Conditional on working and earning positive income (models 2, 4, and 6), improvement in education do not have statistically significant impacts on the probabilities of working for an employee-type of job or working full-time and year-round.

With regards to earnings, the last row of Table 2.4 indicates that the 2SLS estimates of the rates of return to one additional year of schooling are very similar for the two samples: being approximately 18 percent. The 2SLS estimators of the returns to attending college and obtaining any post-secondary degrees are slightly higher for the worker sample: they are estimated to be 78.3 and 91.4 percent for the full sample, and 78.8 and 96.8 percent when restricting to those who reported to have earned positive income in 2013. Coefficients on the monetary returns to education are all statistically significant at least at the 10 percent level.

To put my findings in the context of previous studies, my 2SLS estimators of the returns to an additional year of schooling and college are higher than the returns estimated by previous studies. Using naïve OLS methods, rates of returns to an additional year of schooling are estimated to be around 10 percent during the 2000s (Ge & Yang, 2011; Meng, 2012; Zhang et al., 2005), similar to the world average rate of return to education of roughly 9 percent

(Psacharopoulos & Patrinos, 2018). OLS returns to college are estimated to be in the range of 32-50 percent compared with high school graduates during the 2000s (Giles et al., 2015; Meng, 2012; Zhang et al., 2005).

There are two major reasons why my estimates are higher. First, the 2SLS estimates in Table 2.4 are a weighted average of the returns to education for those who are affected by the instruments (i.e., who attended college due to the expansion, and would not attend college in the absence of the expansion). If the returns to education are not constant among individuals, these estimates may not equal to the OLS estimates of the population (or sample) average returns to education (Angrist and Imbens, 1995). To test this, I use simple OLS regression models with the same sets of controls as in equation (2.3) to estimate the returns to an additional year of schooling, college attendance and post-secondary degree attainment. The naïve OLS estimates of the returns to an additional year of schooling are 7.7 and 9.3 percent for the full and worker sample, very close to those from prior literature. OLS estimates of the returns to attending and completing college are 42.7 and 44.6 for the full sample, and 50.3 and 52.2 percent for the worker sample.

My IV estimates of the returns to college are roughly 56-86 percent higher than the corresponding OLS estimators. This discrepancy is in line with previous studies conducted in the U.S. finding that IV estimates are usually larger than the corresponding OLS estimates (Card 1995, 1999). Also, OLS estimates could be attenuated by measurement errors in education, which will bias the estimates towards zero. Second, to fully capture the effects of the expansion, my sample includes all individuals who have completed lower-secondary school, and thus the comparison group to college attendants or graduates consists of a combination of lower-secondary school graduates, high school dropouts, and high school graduates. Likewise, my treatment groups (those who have ever attended college and obtained any post-secondary degree) include those who go beyond college and have attended or completed graduate school because part of the expansions' education effects is to induce individuals to pursue schooling beyond the baccalaureate level.[24] Previous work on the returns to college largely uses high school graduates as the comparison group and sometimes also excludes those who have graduate degrees. This could explain why my OLS estimates of the returns to college are slightly higher than those estimated by previous work also using OLS regressions.

2.5 Falsification Tests and Robustness Checks

The validity of my method relies on two major assumptions. First, pre-treatment trends in education and labor market outcomes should be parallel for provinces with different expansion intensity. Besides, even in the absence of any differential pre-trends, confounding influences could still happen after the expansion. Thus, we also need to ensure that there are no omitted variables that are correlated with both the treatment intensity and outcomes after the expansion. Below I present results from a series of tests and checks that support the validity of my method.

2.5.1 Falsification Tests

Ideally, I would like to use construct a long panel of province-year level education and earnings outcomes before 1999 to test the parallel pre-trends hypothesis, However, such data are unavailable. Instead, I provide evidence using two alternative methods to show that there are no differential pre-trends in education and labor market outcomes for high and low expansion intensity provinces either across two before-expansion survey years or two groups of before-expansion cohorts.

[24] For the full and worker samples, 2.2 and 2.5 percent of the individuals have any graduate degrees.

For the first test, there are two rounds of the CHIP surveys that were conducted before the college expansion (i.e., CHIP 1988 and CHIP 1995). I append these two pre-expansion surveys to explore whether the growth trajectories in educational attainment, employment and earnings from 1988 to 1995 follow similar temporal trends between the high- and low-expansion-intensity provinces using the following equation:

$$Y_{i,p,y} = c_1 + \alpha \left(\boldsymbol{Survey95_y} * \boldsymbol{HighIntensity_p}\right) + \beta \boldsymbol{Survey95_y} + \gamma \boldsymbol{Province_p} + \theta \boldsymbol{X}_i + \varepsilon_{i,p,y}, \quad (2.4)$$

where $\boldsymbol{Survey95_y}$ indicates whether an individual is from the CHIP1995 survey (it equals zero if an individual is from the CHIP1988 survey), and other variables are defined the same way as in equation (2.1). Similarly, the coefficient of interest is α, which would suggest whether individuals from high expansion intensity provinces enjoy larger increases in education and labor market outcomes from 1988 to 1995 than those from low expansion intensity provinces. Table 2.5 below presents regression results of the DiD estimators of the expansion's effects on education, employment, and earnings.[25]

The DiD estimators for none of the educational and labor market outcomes in Table 2.5 are statistically significant for both the full and worker samples, defined similarly as in my analysis using the CHIP2013 dataset. Moreover, the magnitudes of the coefficients are all very close to zero. These results prove that the growth trajectories in education and labor market outcomes during the pre-expansion period are not systematically different for the high- and low-expansion intensity provinces.

[25] I construct similar sample as that using the CHIP2013 data (i.e., excluding individuals still enrolled in school, retired, disabled, or too old, and restricting to individuals who at least completed lower secondary school), yet the samples are not entirely the same since many survey questions are different for CHIP1988/1995 and CHIP2013.

Table 2.5: Falsification Tests - DiD Estimators of the Expansion's Education and Labor Market Effects Using Two Before-expansion Survey Rounds (i.e., CHIP 1988 & 1995).

	Full Sample *(N=18,435)*		Workers Sample *(N=17,854)*	
	[1]	[2]	[3]	[4]
	raw coefficient	*10 pp. increase in "intensity"*	raw coefficient	*10 pp. increase in "intensity"*
Educational Outcomes:				
Years of schooling completed	-0.04	0.00	-0.05	-0.01
	(0.06)	(0.01)	(0.06)	(0.01)
Attended any post-seconary institution	-0.005	0.000	-0.008	-0.001
	(0.010)	(0.001)	(0.010)	(0.001)
Labor Market Outcomes:				
Probability of working and earning positive income	0.008	0.001	---	---
	(0.005)	(0.001)	---	---
Total annual income	-24	-2	-60	-6
	(78)	(8)	(78)	(8)
(imputed) log(total annual income)	0.019	0.002	-0.032	-0.003
	(0.018)	(0.002)	(0.020)	(0.002)

Data source: Chinese Household Income Project 1988 and 1995.

Notes: Robust standard errors clustered at the province level are in parentheses.

***, **, and * indicate significance at the 0.1, 1, and 5 percent levels.

One caveat of the CHIP1988 and CHIP1995 datasets is that they sample fewer provinces and contain fewer and less detailed information on education and labor market outcomes than the CHIP2013 round. For example, the CHIP1988 and 1995 cover ten and eleven provinces (out of the total of 31 provinces in mainland China) respectively, and do not have any information regarding quantity of labor supplied. Taking these into consideration, results shown in Table 2.5 suggest that there are no pre-existing trends in education, employment and earnings, at least across the provinces covered by the two before-expansion surveys and during 1988 and 1995

For fear that the above test could not capture any trend between 1996 and 1998, or for the other provinces that are not in the CHIP1988/1995 sample but are in the CHIP2013 one, I conduct another falsification test using the CHIP2013 dataset. More specifically, I compare the differences in education and labor market outcomes between two before-expansion cohorts of individuals (those who are scheduled to attend college during 1985-1991 and those during 1992-1998) between high and low expansion intensity provinces. Regression results using equation (2.1) are presented in Table 2.6 below.

Table 2.6: Falsification Tests - DiD Estimators of the Expansion's Education and Labor Market Effects Using Two Before-expansion College Cohorts (i.e., 1985-91 vs. 1992-98).

	Full Sample *(N=3,519)*		Workers Sample *(N=3,178)*	
	[1] raw coefficient	[2] *10 pp. increase in "intensity"*	[3] raw coefficient	[4] *10 pp. increase in "intensity"*
Educational Outcomes:				
Years of schooling completed	-0.1	0.0	-0.2	0.0
	(0.2)	(0.0)	(0.3)	(0.0)
Attended any post-seconary institution	-0.014	-0.001	-0.024	-0.002
	(0.040)	(0.004)	(0.042)	(0.004)
Obtained any postsecondary degree	0.020	0.002	0.010	0.001
	(0.038)	(0.004)	(0.041)	(0.004)
Labor Market Outcomes:				
Probability of working and earning positive income	0.022	0.002	---	---
	(0.025)	(0.003)	---	---
Probability of working for an employee-type of job	0.015	0.001	-0.006	-0.001
	(0.038)	(0.004)	(0.034)	(0.003)
Probability of working full-time-year-round	0.031	0.003	0.011	0.001
	(0.039)	(0.004)	(0.038)	(0.004)
(imputed) log(hourly total income)	0.054	0.005	0.063	0.006
	(0.062)	(0.006)	(0.075)	(0.008)

Data source: Chinese Household Income Project 2013.

Notes: Robust standard errors clustered at the province level are in parentheses.

***, **, and * indicate significance at the 0.1, 1, and 5 percent levels.

For both the full and worker samples, magnitudes of the coefficients on education outcomes in Table 2.6 are all close to zero, and the signs for some coefficients are negative. None of them is statistically significant either. Again, these results provide some suggestive evidence that the observed positive educational effects for highly treated post-expansion cohorts presented in Table 2.2 are not driven by some pre-trends that differentially affected the high- and low-intensity provinces in the absence of the expansion program.

As for labor market outcomes, coefficients on none of the labor market outcomes in Table 2.6 are statistically significant at the 5 percent level. In other words, the 1992-1998 college cohorts (relative to the 1985-1991 college cohorts) in provinces that experienced larger increases in college admission quota ***after*** the expansion show no discernible difference in employment probability, employment stability, and earnings than those from provinces with lower degrees of expansion intensity.

Combined, Table 2.5 and Table 2.6 imply that the divergence in educational outcomes after the expansion is unlikely to be correlated with pre-existing trends in educational outcomes between provinces that experience high and low expansion intensity, but due to the policy change in 1999. Likewise, differences in labor market outcomes between either the two untreated cohorts or the two untreated survey years do not vary across provinces with different expansion intensity, and there are no labor market effects for whom the expansion has no educational effects. This result is reassuring because it suggests that the statistically significant and positive labor market effects presented in Table 2.3 for the more intensively treated individuals should work through the channel of expansion-induced increases in educational attainment. Results from these two sets of falsification tests lend some credibility to the design of my identification strategy.

2.5.2 Robustness Checks

Even though Figure 2.2 shows that there were no systematic inter-provincial differences in the changes of college admission ratio before the expansion, and the falsification tests above suggest that there were no pre-trends in educational and labor market outcomes across provinces with different degrees of expansion intensity, one could still argue that potential confounding factors that simultaneously affect the province-cohort level expansion intensity and the outcomes emerged just at or after the time of the expansion. In other words, there might be some province-level time-varying variables that are correlated with both the expansion intensity measure and the outcomes in 2013.

For this concern, I test for a series of province-year level covariates and concurrent events that might correlate with both the treatment and potential outcomes, and add them as covariates into equation (2.1), thus using only the exogenous part of variation in the intensity measure to identify the expansion's causal effects on education and labor market outcomes.

In practice, I first test whether there are province-level confounding factors prior to the expansion that determine the post-expansion intensity measure. Then, I explore whether there are time-varying province-level confounding influences after the expansion was first implemented, and add them interacted with post-expansion year indicators as covariates in equation (2.1). The confounding factors I test for include: various measures of province-year level annual growth rate of per capita GDP, various measures of province-year level unemployment rate, province level number of higher education institutions in 1998 (i.e., before the expansion), and dummies for whether the province is more or less affected by two influential concurrent reforms (China's accession into the World Trade Organization in late 2001 and the reforms of state-owned enterprises that span from the early 1990s to early 2000s). I explain each of these variables in more details in Appendix B.

In addition, I also test for the existence of agglomeration effect. In particular for my sample, Beijing and Shanghai are two provincial level autonomous cities that have the best (higher) education and economic conditions in China. As Panel A of Figure 2.2 illustrates, Beijing and Shanghai are outliers for college admission rate both before and after the expansion. For example, at the baseline level in 1998, 32 and 29 out of a hundred 18-year-olds in Beijing and Shanghai could be admitted into colleges, compared with an average of less than 6 students across all provinces in China. Consequently, the educational effects from a 10-percentage-points increase in the college admission rates for Beijing and Shanghai could be different from that for other provinces. Moreover, Beijing and Shanghai have the largest and growing proportion of skill-intensive industries, suggesting that they could have disproportionately higher and increasing demand for high-skill labor than other provinces in China. Lastly, these two metropolitan cities also have the largest proportion of workers who work there without local "hukou." For these reasons, I replicate all the analysis presented in section 4 excluding individuals whose "CEE province" is Beijing or Shanghai, which account for 7.1 and 7.0 percent of the full and worker samples respectively.

Results from these exercises show that for the first year of the expansion (i.e., 1999), the expansion "intensity" is not determined by province-level covariates, including per capita GDP growth rate from 1998 to 1999, lagged per capita GDP growth rate 4 years later (i.e., from 2002 to 2001), unemployment rate in 1999 or 1998, nor by any province-level pre-trends in per-capita GDP growth rate or unemployment rate from 1995 to 1998. The only statistically significant determinant of "intensity" is the province-level number of higher education institutions in 1998, and I further include this variable as a control in equation (2.1). However, there are some indication that as the expansion continued, correlations between province-year level “intensity” and the above-mentioned covariates: for example, both or one of current year’s per capita GDP growth rate and unemployment rate have statistically significant effects on the “intensity” measure for some of the post-expansion years starting from 2003. Therefore, as a robustness check, I use equation (2.5) that further adds potential confounding time-varying province-level covariates for post-expansion years to my baseline model shown in equation (2.1).

$$Y_{i,p,c} = c_1 + \alpha\left(\boldsymbol{Post_c} * \boldsymbol{Intensity_{p,c}}\right) + \beta\boldsymbol{Cohort}_c + \gamma\boldsymbol{Province}_p + \theta\boldsymbol{X}_i + \delta\left(\boldsymbol{Post}_c * \boldsymbol{Covariates_{p,c}}\right) + \varepsilon_{i,p,c}. \tag{2.5}$$

Table 2.7 below replicates Table 2.2 and Table 2.3 on the expansion’s education and labor market effects excluding Beijing and Shanghai, and further with additional time-varying province-year level covariates, for the full and worker samples respectively.

For the three educational outcomes, magnitudes of the coefficients become much larger for both samples after excluding the two largest metropolitan cities in China. The expansion’s educational effects become even more pronounced after further controlling for potential province-cohort level post-expansion cofounding factors. For example, for the full analysis sample, a 10-percentage-point increase in expansion intensity leads to approximately 0.6 more years of schooling, 10.4 and 9.9 percentage points higher probabilities of attending any college and obtaining any post-secondary degrees (second column under model 1 of Table 2.7), compared with increases of 0.3 more years, and 7.2 and 6.1 percentage points using all the provinces (column 2 of Table 2.2).

The coefficients become even larger after controlling for potential confounding covariates: a 10-percentage-point increase in expansion intensity leads to approximately 0.9 more years of schooling, and this magnitude triples the corresponding education effect estimated using the baseline model; a 10-percentage-point increase in intensity also leads to approximately 14.0 and 13.8 percentage points higher probabilities of attending college and obtaining a post-secondary degree – roughly double the corresponding effects estimated using the baseline model.

Table 2.7: Robustness Checks - Excluding Beijing and Shanghai, with and without Province Covariates.

	Full Sample *(N=7,782)*						Workers Sample *(N=6,610)*					
	[1]			[2]			[3]			[4]		
	raw coef.	10 pp. increase in intensity		raw coef.	10 pp. increase in intensity		raw coef.	10 pp. increase in intensity		raw coef.	10 pp. increase in intensity	
Educational Outcomes:												
Years of schooling completed	6.1	0.6	***	8.8	0.9	***	5.8	0.6	***	8.8	0.9	***
	(1.6)	(0.2)		(1.9)	(0.2)		(1.7)	(0.2)		(2.1)	(0.2)	
Attended any post-seconary institutions	1.040	0.104	***	1.404	0.140	***	1.032	0.103	***	1.475	0.147	***
	(0.257)	(0.026)		(0.315)	(0.031)		(0.289)	(0.029)		(0.354)	(0.035)	
Obtained any postsecondary degrees	0.993	0.099	***	1.382	0.138	***	0.956	0.096	***	1.413	0.141	***
	(0.250)	(0.025)		(0.306)	(0.031)		(0.282)	(0.028)		(0.346)	(0.035)	
Labor Market Outcomes:												
Probability of working and earning positive income	0.249	0.025		0.123	0.012		---	---		---	---	
	(0.199)	(0.020)		(0.244)	(0.024)		---	---		---	---	
Probability of working for an employee-type of job	0.480	0.048	~	0.424	0.042		0.235	0.023		0.319	0.032	
	(0.257)	(0.026)		(0.315)	(0.031)		(0.227)	(0.023)		(0.279)	(0.028)	
Probability of working full-time-year-round	0.201	0.020		0.050	0.005		-0.002	0.000		-0.061	-0.006	
	(0.274)	(0.027)		(0.336)	(0.034)		(0.273)	(0.027)		(0.335)	(0.034)	
Probability of working under a long-term employment contract	0.188	0.019		-0.151	-0.015		0.054	0.005		-0.256	-0.026	
	(0.300)	(0.030)		(0.366)	(0.037)		(0.335)	(0.033)		(0.409)	(0.041)	
(imputed) log(hourly total income)*	0.107	0.011		0.005	0.001		0.152	0.015		0.122	0.012	
	(0.387)	(0.039)		(0.474)	(0.047)		(0.516)	(0.052)		(0.632)	(0.063)	
Excl. Beijing & Shanghai	YES			YES			YES			YES		
Incl. post-expansion province-cohort level covariates**	NO			YES			NO			YES		

Data source: Chinese Household Income Project 2013.

Notes: Total income includes income from all jobs that include salary, bonus, compensation, and in-kind subsidies transferred to the Chinese yuan values; Province-cohort level covariates for post-expansion years include: annual per-capita GDP growth rate, annual unemployment rate, dummy variables indicating whether a province has an above-national-mean proportion of export and import industries for a certain year and thus more affected by China's WTO accession, and the number of higher education institutions in 1998.

Robust standard errors clustered at the province level are in parentheses.

***, **, *, and ~ indicate significance at the 0.1, 1, 5, and 10 percent levels.

This pattern of results after excluding Beijing and Shanghai and including additional controls is very similar for the worker sample. The magnitudes of the educational effects become one to two times larger than those estimated using the baseline model. Coefficients on all the three educational outcomes remain statistically significant at the 0.1 percent level.

On the contrary, with regards to labor market outcomes, the coefficients on the expansion's labor market effects generally become smaller after excluding Beijing and Shanghai, and further shrink after controlling for post-expansion province-year level covariates. For example, for the full sample the expansion's earnings effects become less than one fifth of that estimated using the baseline model, and further reduces to be very close to zero after controlling for additional covariates (last row under the second column of models 1 and 2 of Table 2.7).

I observe the same pattern for the worker sample: a 10-percentage-point increase in expansion intensity leads to approximately 1.5 and 1.2 percent increases in the hourly wage when excluding workers from Beijing and Shanghai, and further including additional covariates respectively, compared with a 5.3 percent increase using the baseline model (last row of column 4 in Table 2.3). Moreover, coefficients on none of the labor market outcomes remain statistically significant at the 5 percent level after excluding Beijing and Shanghai, or further adding the covariates.

Results from comparing the estimated effects when including and excluding Beijing and Shanghai imply that the expansion has even larger positive educational effects for provinces with lower access to higher education before the expansion, but weaker effects on earnings for less developed and human capital intensive provinces. Since Beijing and Shanghai experienced the most substantial growth in the high-tech and other human capital intensive industries during the 2000s, these two cities could have much higher increases in the demand for college-educated labor than other provinces in China. The discrepancy in estimated earnings effects between including and excluding the two largest cities suggest that the estimated positive labor market outcomes presented in Table 2.3 are partly driven by the subgroup of individuals from big cities. Results from the earnings effects after excluding these individuals might imply that the skill price of college-educated labor could be depressed if the increase in demand for them cannot keep pace with the surge in supply of them in the local labor market.

Results from comparing the estimated effects with and without the covariates show that the estimated expansion's educational effects are downwardly biased if not including these covariates. This is understandable since all the four covariates are positively correlated with expansion intensity, and provinces with higher GDP growth rate, more higher education institutions in 1998, and highly affected by the two concurrent events, on average, have higher educational attainment to begin with. Consequently, the improvement from older unexposed cohorts to younger exposed cohorts could be smaller than provinces with lower level of educational attainment before the expansion. In contrast, most covariates are positively related to labor market outcomes, and thus producing upward biases in the estimated employment and earnings effects of the expansion when no such covariates are included.

Results from the above robustness checks illustrate that the expansion has substantial and positive effects on education, and the magnitudes of these effects double to triple after excluding individuals from Beijing or Shanghai, and further control for potential confounding influences after the expansion. However, most of the expansion's positive labor market effects estimated using the baseline model shrink and become statistically insignificant after excluding the two largest metropolitan cities in China and further controlling for additional covariates. The latter finding implies that the positive labor market outcomes observed in Table 2.3 mainly concentrate in large and more developed areas where there are more human capital intensive industries and a rising demand for college-educated labor.

2.5.3 Alternative Channels for the Expansion's Effects

Lastly, I address some concerns for the interpretation of the estimated expansion's effects, including potential biases due to selective migration and potential confounding influence from changes in the quality of education.

Selective Migration. Students could migrate out of their "hukou" province to take the CEE, and this assertive migration could bias the expansion's educational effects. Also, graduates could migrate to other provinces to work, and this selective migration may bias the estimates of the expansion's earnings effects. As explained previously, migration across provinces is not a serious problem for the CHIP 2013 sample since it tends to sample communities that are occupied by local households: only 2.4 percent out of the full analysis sample took the CEE or reside/work in a different province in 2013 than their "hukou" province, and the proportion is 2.5 percent for the worker sample.[26]

However, I re-run all the models excluding those individuals who ever changed their "hukou" province to take the CEE, or who work in a province other than their "hukou" province in 2013. The major take-away is that there are minimal changes after excluding these small proportion of individuals, and there is no pattern showing that migrants in the CHIP 2013 sample are systematically different in any of the characteristics presented in Table 2.1 from non-migrants.

[26] According to surveys of recent college graduates, the majority of students attend post-secondary institutions in their home province, and work in the province where they attend college (Peking University School of Educaton, 2011).

Changes in Education Quality. One concern with the expansion is that it might affect both the quantity and quality of education, and as a result, the estimated wage effects might not only come from increased years of schooling (and the attainment of degrees), but also from any changes in education quality due to the expansion. To explore whether this is the case, I use two commonly used measures of education quality, student-faculty ratio and per-student expenditure (Barr & Turner, 2013; Bound & Turner, 2007; Chakrabarti, Gorton, & Lovenheim, 2017), to depict whether there have been changes in the quality of education before and after the expansion.

Table A5 in Appendix A shows the number of regular HEIs, total undergraduates enrollment, and average student-faculty ratio in such institutions during 1995-2018. Number of regular HEIs more than doubled, and total enrollment grew by nearly tenfold during this period, with the most radical increases happening during the years between 1999 and 2007. Particularly, the last column in Table A5 shows a trend of decreasing resources per student measured using student-faculty ratio: this ratio nearly doubled as average enrollment per institution more than tripled.

Using another common proxy for education quality (i.e., expenditure per student), Figure A1 in Appendix A illustrates the trends in total funding per full-time-equivalent (FTE) student, decomposed into government appropriations, tuition and fees, and other sources (including donation, revenue from cooperating with industry/companies and college-run enterprises, etc.). We can see that after the expansion, per FTE student funding (i.e., resources available to institutions to educate each student) has been at least at the same or slightly higher level compared with that in 1998 (except for a minor decrease for the year of 2013). In fact, there is a six percent increase in per FTE student funding in the first year of the expansion, and larger increases starting from 2007. However, such increases were mainly driven by increases in tuition and fees, which offset the decreasing per-FTE-student funding from government appropriations during 1999-2005.

These descriptive statistics imply that the average level of education quality might decrease as the average level of education increases during the expansion era. This could bias the estimates of the expansion's earnings effects downwardly. It is noteworthy that the changes in education quality are likely to be different for HEIs of different tiers and types. Without data at the institutional, or at least the tier level, we cannot rigorously assess the expansion's effects on the quality of college education.

2.6 Discussion

There are two major limitations for this analysis, and more broadly, for using one cross-section of data to estimate the treatment effects of large-scale social programs under the partial-equilibrium framework. I elaborate on both limitations in this section.

2.6.1 Limitation of Using One Cross-section of Data

One caveat of this analysis is that the earnings data are observed at one time in 2013, which is 14 years after the expansion started and 2-11 years after the post-expansion college graduates first entered the labor market (when they are between 23 to 33 years old). For one thing, Jackson, Johnson, and Persico (2016) compare earnings returns estimated using panel and cross-sectional data and find that using a single year of earnings data yields point estimates one-half to two-thirds of and standard errors four times as large as those estimated using multiple years of earnings data. For the other thing, the exposed are still relatively young, and their earnings might not reflect their permanent earnings (for example, their labor market outcomes might take off and/or become more stable at older ages). Jackson, Johnson, and Persico show that returns to education are lower when earnings are observed at younger ages than that using earnings at older ages, even after controlling for a quadratic term of age. Haider and Solon (2006) find similar results, and conclude that using earnings data from one single year and/or at young ages will under-estimate the earnings returns. For the purpose of analyzing the earnings effects of China's college expansion, the CHIP 2013 dataset I am using is so far the best possible one that includes the most recent round of earnings data for post-expansion cohorts, as well as contains comprehensive information on individuals' demographic, educational, and labor market conditions. In next chapter, I will use multiple-rounds of cross-sectional data from the State Bureau of Statistics' Urban Household Survey (with most recent round of earnings data from 2009, and less detailed demographic, education, and labor market information) to partly deal with the limitation of using one single round of data.

2.6.2 Limitation of Estimating Treatment Effects When General-Equilibrium Effects Present

The quasi-experimental methods used in this analysis are under the assumptions that there are no spill-overs to those whose behaviors are unaffected by the treatment (i.e., the always-takers and never-takers under the LATE framework). These methods work best in micro-interventions when all the other conditions can be held constant as the "treated" individuals change their schooling decisions However, large-scale policy changes like education expansions, tuition reforms and schooling law modifications tend to alter labor market dynamics and affect people who are not directly treated, leading to general equilibrium effects. Particularly with this case, the surge in college-educated labor due to the expansion could depress the equilibrium price for college labor, which would in turn affects earnings of older unexposed cohorts of college graduates even though they are not treated by the expansion.

Moreover, there is no pure control group post-expansion in this case since every province increases college admission spots to some extent for each of the years after the expansion. This is also the case for other large-scale policy changes/programs (Duflo, 2001; Lemieux & Card, 2001).

Except for potential changes in the equilibrium skill prices, there could also be behavioral responses. For example, younger students might also respond to the depressed college premiums by altering their college attendance decisions, and firms might re-allocate resources between human capital and physical capital, and/or adopt some endogenous technological changes. These responses could violate the exclusion restriction assumption under the IV method, that is, the instrument (i.e., college expansion) affects potential outcomes only through the endogenous variable (i.e., education).

In sum, even using credible sources of exogenous shocks, the estimated treatment effects are no longer the partial equilibrium returns to education, but at most an aggregate effect of the policy that captures both the educational returns and other spillover effects. However, I cannot parse apart these two components or delve into the exact magnitude of the general-equilibrium effects with only one cross-section of earnings data. More specifically, the general equilibrium effects of the expansion program depend on the elasticity of the demand for college graduates, the elasticity of substitution across cohorts/age groups, students' sensitivity of educational choices to perceived changes in the returns to college, and the extent to which China has a segregated local labor markets or an integrated national labor market.

Chapter 3 of my dissertation can somewhat address this caveat. I use multiple repeated cross-sections of data to explore how the wage structure (especially the college-High school wage gap) evolved after the expansion. Using multiple cross-sections of earnings data, I could identify shifts in the aggregate and cohort-specific supply of college-educated labor, potential shift in the demand for college-educated labor, and changes in the quality composition of the college- vs. high school-educated labor. This third essay allows for changes in the supply and demand of different kinds of labor over time, and allows relative skill prices to endogenously adjust to such changes. By doing so, I can therefore decompose simultaneously changing labor market conditions and parse apart different kinds of effects in general equilibriums.

2.7 Summary

In this chapter, I provide new evidence on the enduring question of whether and to what extent attending and completing college affects an individual's subsequent labor market outcomes. I take advantage of a large-scale reform - the college expansion program in China - to isolate exogenous changes in the probability of attending college that are uncorrelated with unobserved determinants of an individual's labor market outcomes. I use this expansion-induced increases in educational attainment to estimate the causal returns to an additional year of schooling, college attendance, and postsecondary degree attainment.

I link education and earnings information of a large cross-section of individuals born between 1959 and 1990 (thus attending college between 1977 and 2008) from the 2013 CHIP with province-cohort level administrative records on the number of college admission spots, and combine differences across cohorts induced by the timing of the college expansion with differences in the additional admission quotas allocated across provinces and cohort years to estimate the effects of the expansion on education and earning. Results suggest that the expansion substantially increases individuals' educational attainment, including completed years of schooling, and the probabilities of attaining college and obtaining any postsecondary degrees. The expansion also increases individuals' probability of working and their earnings in 2013.

However, the expansion's earnings effects are less robust to the exclusion of the two largest metropolitan cities in China and the inclusion of additional province-level time-varying covariates after the expansion. Combining the expansion's education and earnings effects implies estimates of the economic returns to college attendance and post-secondary degree attainment ranging from 78-79 and 79-97 percent respectively. Results from my analysis reinforce that improved access to college can substantially improve students' educational attainment, and eventually benefit their life outcomes.

Chapter 3: Evolution of the College Earnings Premium for Younger and Older Workers during the 2000s in China

3.1 Introduction

As China shifted from a centrally planned economy to a market-oriented one since its major economic transformation in 1978, the earnings differentials across workers with different educational attainment substantially widened. For example, ordinary least squares (OLS) estimates of the returns to 4-year college (as compared to high school only) rose from about 13 percent in the late 1980s to an average of 23 percent during the 1990s, and peaked at approximately 40 percent in 2000 and 2001 (Zhang et al., 2005). There are mixed findings regarding the trends in the college premium after 2001: some studies show that the college earnings premium declined during the 2000s (Appleton, Song, & Xia, 2014; Yao, Xin, & Zhang, 2013), while others find the premium to kept increasing (Carnoy et al., 2012; Hu & Hibel, 2014; Yang & Gao, 2018) or stayed relatively constant (Gao & Smyth, 2015) during the 2000s.

In this chapter, I intend to depict the evolution of college earnings premium during the 2002 and 2009, a period when China experienced an increasing degree of openness and trade due to its accession into the World Trade Organization (WTO) in 2001, which presumably would increase the demand for high-skill, college-educated labor (Pack & Nelson, 1999; Noornackhsh, Paloni, & Youssef, 2001). Moreover, this is a period when the first cohort(s) of college graduates affected by China's large-scale higher education expansion program entered the labor market. The expansion was initiated in 1999, and enrollment of first-year-first-time undergraduate students saw the most substantial increases during the first three years of the expansion. The first after-expansion cohorts of 3- and 4-year college graduates would enter the labor market starting in 2002 and 2003, and the sampled period would capture more and more after-expansion cohorts during later years. It is thus interesting to document the labor market dynamics during this time, and the evolution of education-related earnings differentials resulting from the interplay between changes in both the demand and supply of college-educated labor.

The rest of this chapter is organized as follows: in Section 2, using multiple repeated cross-sections of data from the annual Urban Household Survey, I depict the trends in college premium during 2002-2009 for individuals of different age groups. In Section 3, I set up a theoretical model of aggregate labor supplies with imperfect substitution across age groups to explain the observed pattern presented in the previous section. Section 4 introduces the data and estimation procedure used to specify model parameters. In Section 5, I present results from decomposing the 4-year college premium into the cohort-specific and aggregate relative supply, and aggregate relative demand factors. Section 6 concludes and discusses the potential policy implications of my study.

3.2 Evolution of College Earnings Premium during 2002-2009

Using eight repeated cross-sections of earnings data from China's annual Urban Household Survey (UHS) over the period from 2002 to 2009, I find that the BA-HS earnings gaps do not rise and fall in a parallel manner for workers of different age groups. More specifically, I calculate the mean logarithm of annual job income for workers of different age groups who have reported positive job income by year and separately for whose highest education levels are 4-year college (BA-educated) and high school (HS-educated) respectively. Then I compute the gap in the mean logarithm of annual job income between BA- and HS-educated labor within each year-age group cell (see Table A6 in Appendix A for education level-age group-year level mean income during 2002-2009). Figure 3.1 below depicts the trends in the BA-HS gap for workers in selected age groups who earn positive income.

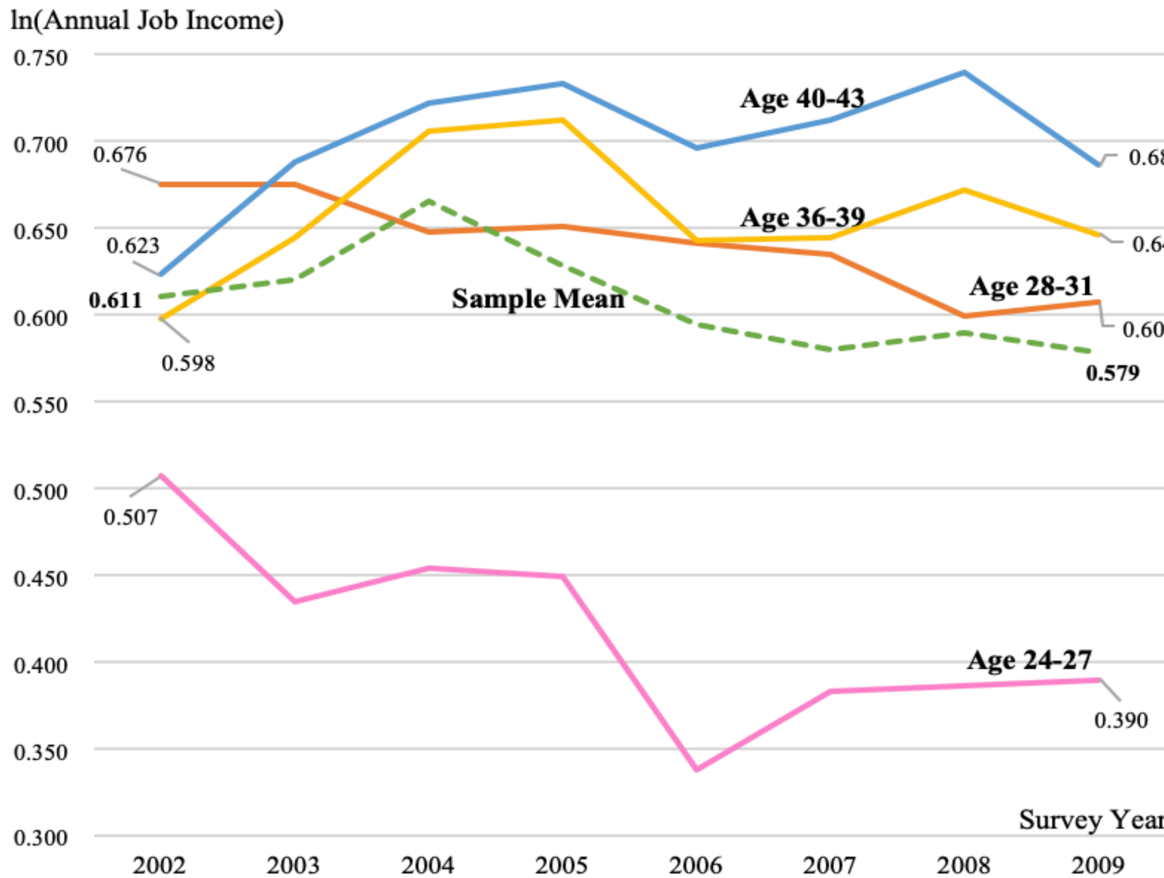


Figure 3.1: BA-HS Earnings Gap by Year for the Worker Sample and by Selected Age Groups.

Data source: Urban Household Survey, 2002-2009.

Notes: Annual incomes are discounted to the 2002 Chinese currency (¥) value using annual nationwide CPIs in urban areas obtained from the China Statistical Yearbooks, 2002-2009.

The green dotted line in Figure 3.1 illustrates how BA premium changes during 2002-2009 for all the workers. It increases during 2002-2004, and dropped afterwards (with a re-bounce in 2008). By 2009, the BA premium is approximately 3 percentage points (or 5 percent) lower than its 2002 level. Note that the period under this study corresponds to a substantial surge in the supply of BA-educated labor due to China's higher education expansion that was initially in 1999. As shown by Figure 3.2, the number of graduates from 4-year colleges substantially increased starting from 2003. This first after-expansion cohort of 4-year college would enter the labor market in 2003, and their earnings would be captured by my analysis starting from 2005 (when they are at least 24 years old and worked for at least an entire year). By 2007 the three after-expansion cohorts that experienced the most radical college enrollment expansion could have all entered the labor market. Therefore, the trend illustrated by the green dotted line in Figure 3.1 could be related to this increase in the supply of BA-educated labor.

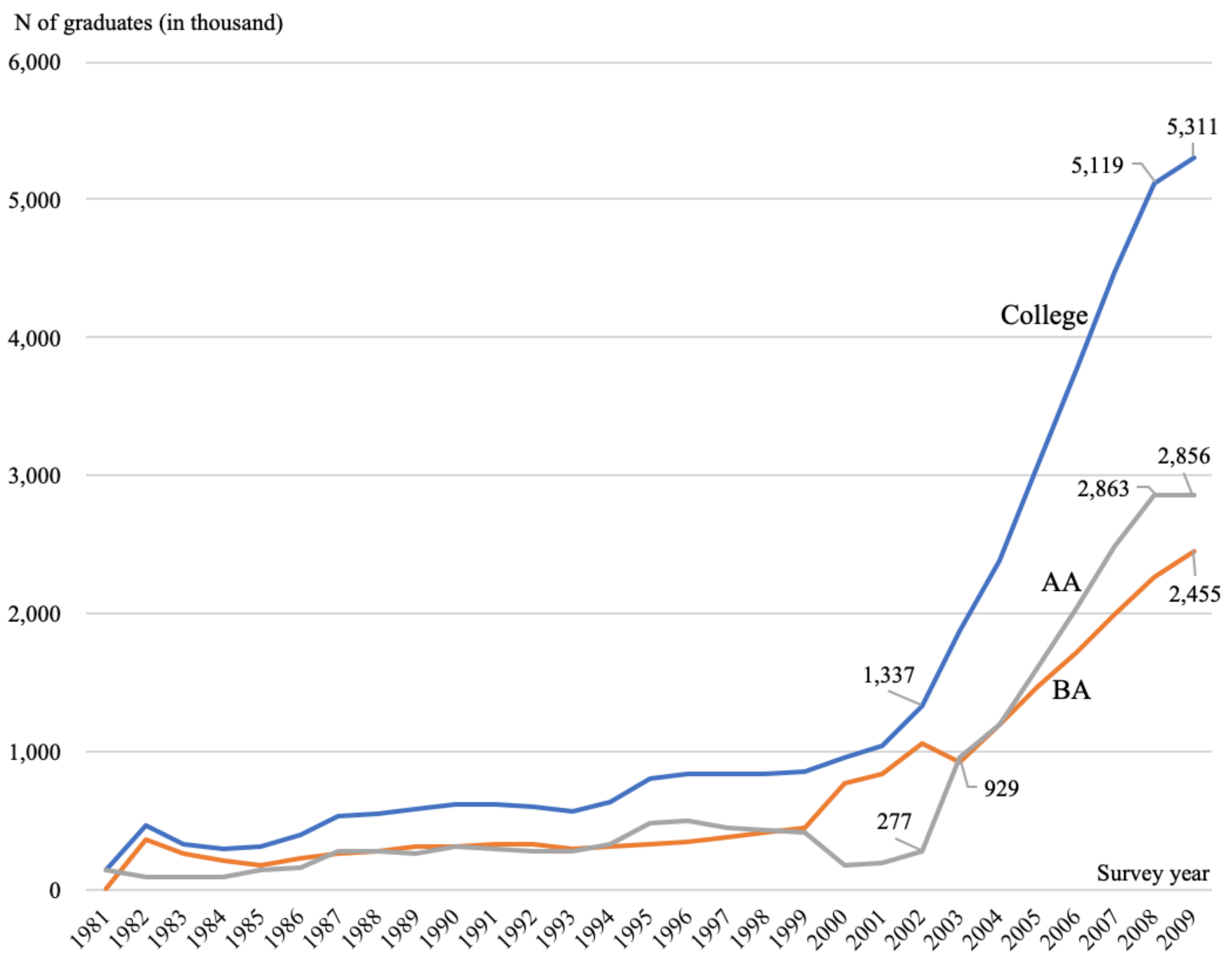


Figure 3.2: Number of Graduates from 3- and 4-year Regular Higher Education Institutions by Year.

Data source: Educational Statistics Yearbook of China, 1981-2009.

However, this sample mean BA premium averaged across all ages masks some distinct features of age-group specific BA premiums. In Figure 3.1, I also present trends of the BA-HS gaps specifically for two old (ages 36-39 and 40-43) and young age groups (ages 24-27 and 28-31) respectively. The trajectories of the evolution of the BA-HS gaps are quite parallel for the two old age groups: they both increase until 2005, and decline in 2006; the gaps were widened again gradually in 2006 and more substantially in 2007 before another dip in 2009. By 2009, the BA-HS income gaps are approximately at the same level as in 2006, but still higher than their 2002 levels (10 percent higher for the 40-43 age group and 8 percent for the 36-39). On the other hand, the younger groups experienced quite divergent trends: the BA-HS income gap have been gradually dropping for those aged 36-39 approximately all the way through this period (with tiny increases in 2005 and 2009), ending 10 percent lower in 2009 than its baseline level in 2002; the decreases are much more substantial for the youngest age group (ages 24-27): it experienced two big dips in 2003 and 2006, and the gaps are mostly stagnated during the other years. By 2009, the BA-HS income gap, or the BA premium, is 12 percentage points, or 23 percent, lower than it is in 2002. We see similar patterns using the level form of earnings. I present trends in earnings by age group for HS- and college-educated labor over the same period separately in Figure A2 and their corresponding earnings gaps in Figure A3 in Appendix A.

Under the hypothesis of perfect substitution across different age groups within the same education level, we would expect to see parallel trends for workers in different ages, with their earnings gaps fully accounted for by age effects. This framework fail to explain the patterns observed in Figure 3.1. However, I will show in this essay that relaxing the hypothesis of perfect substitution and incorporating imperfect substitution across workers of the same education level but in different age groups could provide a unified explanation for the observed divergent trends in earnings growth and consequently the BA-HS wage gaps for the younger and older workers.

My study is in the tradition of Card & Lemieux (2001) who aim to explain the rising college-HS wage gaps during 1970s-1990s in the U.S., UK, and Canada. They propose a model that allows for imperfect substitution between younger and older workers with the same education level, and use data from U.S., UK, and Canada to identify the model.[27] Their results show similar inter-cohort trends with steadily rising educational attainment for cohorts born up to 1950, and relative stagnation for the baby-boom cohorts. Moreover, the derived parameters using data from the three countries are comparably similar, suggesting that the model is robust to different educational systems. The authors thus conclude that the observed wage patterns in the three countries can be largely explained by: (1) cohort-specific relative supply of college-educated labor, and (2) steady rises in the relative productivity of college-educated workers (i.e., due to skill-biased technical change). These findings suggest that the cohorts who experienced substantial changes in the cohort-wide average educational attainment either due to demographic reasons or policy changes could have additional cohort effects in addition to the effects equally spread to all cohorts of the same education group.

My study is also related to the larger literature on the evolution of income inequality over time, especially inequality by education. Empirical studies of the U.S. labor market put forth the following explanations for the changes in the college-HS wage gap (or college premium): (1) changes in the relative supply due to college enrollment, women's labor force participation, or immigration (Acemoglu, Autor, & Lyle, 2004; Autor, Katz & Kearny, 2008; Edlund & Kopczuk,

[27] For the U.S., they use the 1960 Census and the March Current Population Surveys (CPS) from 1970-1997 to study men born in 1959-1996. For UK, they use the 1974-1996 General Household Surveys, and the 1981, 1985, 1991, & 1996 Censuses for Canada.

2009; Kopczuk, Saez, & Song, 2010); (2) changes in the relative demand for college-educated labor due to trade, and skill-biased technological change (Acemoglu, 1998; Acemoglu & Restrepo, 2018; Autor, Katz, & Krueger, 1998; Autor & Murnane, 2003, 2015; Berman, Bound, & Griliches, 1994; Card & DiNardo, 2002; Goldin & Katz, 2007b; Lemieux, 2006; Spitz-Oener, 2006); (3) changes in the relative quality of college education and/or the relative ability of college graduates (Blau & Kahn, 2005; Carneiro & Lee, 2011); (4) institutional changes, such as changes in minimum wage or unionization (Autor, Manning, & Smith, 2010; Brown, 1999; Card, 2001; DiNardo, Fortin, & Lemieux, 1996); and combinations of multiple above-mentioned factors (Autor, 2014; Beaudry, Green, & Sand, 2014; Juhn, Murphy, & Pierce, 1993; Katz & Autor, 1999; Katz & Murphy, 1992; Piketty & Saez, 2003).

3.3 Theoretical Framework

I adapt the model from Card and Lemieux (2001) to explain how the wage gaps between two types of labor over time are affected by the aggregate and age specific supplies of these labors. I first focus on the wage gaps between individuals whose highest educational levels are high school and 4-year college (i.e., the BA-HS wage gaps) because 3-year colleges in China (that grant associate's degrees) went through substantial institutional reforms and re-structuring during the past decades. As a result, the productivity and/or ability of 3-year-college-educated labor could vary a lot across cohorts under study by this study.

First, the model assumes that there are only two types of labor supplied: high school (HS)-educated labor (L_{HS}) and 4-year-college-educated labor (L_{BA}), following Card & Lemieux (2001) and further assumes that there is imperfect substitution across labor with different ages within the same educational level, then the sub-aggregates of HS-educated labor ($L_{HS,t}$) and BA-educated labor ($L_{BA,t}$) at year t can be written as:

$$L_{HS,t} = [\ \sum_j \left(\alpha_j l_{HS,j,t}^{\eta}\right)\ ]^{1/\eta}, \tag{3.1}$$

and

$$L_{BA,t} = [\ \sum_j \left(\beta_j l_{BA,j,t}^{\eta}\right)\ ]^{1/\eta}, \tag{3.2}$$

where $l_{HS,j,t}$ and $l_{BA,j,t}$ are age-specific supplies of HS- and BA-educated labor from age group j at year t respectively. $\eta = 1 - \frac{1}{\sigma_A}$, and σ_A is the elasticity of substitution between labor in different age groups *j* within the same education group. Note that when $\sigma_A = +\infty$ (i.e., $\eta = 1$), that is, when individuals of different ages with the same education level are perfect substitutes for each other, the sub-aggregates of HS- and BA-educated labor are just simple (or weighted) sums of the quantities of labor supplied by all age groups within the same education levels (i.e., $l_{HS,j,t}$ and $l_{BA,j,t}$) respectively.[28] α_j and β_j are a vector of relative efficiency parameters across *age groups* for HS- and BA-educated labor respectively (assumed to be the same across years).

Aggregate output Y_t at year t depends on the aggregate level of labor $\boldsymbol{L}_t$ and capital $\boldsymbol{K}_t$:

$$Y_t = F(\boldsymbol{L}_t, \boldsymbol{K}_t). \tag{3.3}$$

The aggregate labor supply $\boldsymbol{L}_t$ depends on the supplies of two types of labor. Assuming the aggregate production function satisfies constant elasticity of substitution (CES), then

$$\boldsymbol{L}_t = (\theta_{HS,t}\, L_{HS,t}^{\rho} + \theta_{BA,t}\, L_{BA,t}^{\rho})^{\frac{1}{\rho}}, \tag{3.4}$$

where $\rho = 1 - \frac{1}{\sigma_E}$, and σ_E is the elasticity of substitution between the two types of labor (i.e., $L_{HS,t}$ and $L_{BA,t}$). $\theta_{HS,t}$ and $\theta_{BA,t}$ are the productivity parameters or technology efficiency parameters for labor of the two *education levels* at year t respectively.

[28] The "simple" sum assumes that labor in different ages supply the same efficiency unit of labor, and the "weighted" sum allows labor in different ages to supply different efficiency units of labor (with the weights often determined by the relative mean wage of the age group).

In a competitive labor market and in equilibrium, a HS-educated worker in age group j at year t gets paid his/her marginal productivity of labor so that:

$$\omega_{HS,j,t} = \frac{\partial Y_t}{\partial l_{HS,j,t}} = \frac{\partial Y_t}{\partial L_{HS,t}} \times \frac{\partial L_{HS,t}}{\partial l_{HS,j,t}} = \theta_{HS,t} L_{HS,t}^{\rho-\eta} (\theta_{HS,t} L_{HS,t}^{\rho} + \theta_{BA,t} L_{BA,t}^{\rho})^{\frac{1-\rho}{\rho}} \times \alpha_j l_{HS,j,t}^{\eta-1}. \quad (3.5)$$

Similarly, a BA-educated worker in age group j at year t gets paid his/her marginal productivity of labor so that:

$$\omega_{BA,j,t} = \frac{\partial Y_t}{\partial l_{BA,j,t}} = \frac{\partial Y_t}{\partial L_{BA,t}} \times \frac{\partial L_{BA,t}}{\partial l_{BA,j,t}} = \theta_{BA,t} L_{BA,t}^{\rho-\eta} (\theta_{HS,t} L_{HS,t}^{\rho} + \theta_{BA,t} L_{BA,t}^{\rho})^{\frac{1-\rho}{\rho}} \times \beta_j l_{BA,j,t}^{\eta-1}. \quad (3.6)$$

Efficient utilization of these two types of workers imply that their relative wages reflect their relative marginal product. Under this condition, age-group specific gap between wages measured in natural logarithm of the two types of labor at year t can be written as:

$$r_{jt} \equiv \ln(\frac{\omega_{j,t}^{BA}}{\omega_{j,t}^{HS}}) = \ln\left(\frac{\theta_{BA,t}}{\theta_{HS,t}}\right) + \ln\left(\frac{\beta_j}{\alpha_j}\right) + \left(\frac{1}{\sigma_A} - \frac{1}{\sigma_E}\right) \ln\left(\frac{L_{BA,t}}{L_{HS,t}}\right) - \frac{1}{\sigma_A} \ln\left(\frac{l_{BA,j,t}}{l_{HS,j,t}}\right) + e_{jt}, \quad (3.7)$$

where e_{jt} is the error term that represents other sources of variation in the age specific wage gaps.

Equation (3.7) implies that the BA-HS wage gap (or the college premium) for individuals in age group j at year t depends on both the aggregate supplies of BA- and HS-educated labor in the labor market at time t (i.e., $L_{HS,t}$ and $L_{BA,t}$), and the corresponding age specific supplies of respective labor within his/her own age cohorts (i.e., $l_{HS,j,t}$ and $l_{BA,j,t}$).

3.4 Data and Method

The aim of this study is to identify all the parameters in equation (3.7), and the main focus is on those before the age group specific (i.e., $-\frac{1}{\sigma_A}$) and aggregate (i.e., $\frac{1}{\sigma_A} - \frac{1}{\sigma_E}$) supplies: the former could help explain the fall in college premiums particularly accrued to younger cohorts, and the latter could capture the general-equilibrium effects of an rise in college attainment (and thus the supply of college-educated labor) on all cohorts in the labor market. In the following paragraphs, I elaborate on how I use data from China to empirically identify the model presented in equation (3.7), and thus to explain the evolution of the college wage premiums during the 2000s.

3.4.1 Data

I use data from the Urban Household Survey (UHS) conducted annually by China's National Bureau of Statistics (NBS). NBS adopts a probabilistic and stratified multi-stage sampling method to obtain a nationally representative sample for each round (Meng, 2012). The sample size is large: the full sample contains approximately 100,000 individuals for each round. Compared to other household surveys including the CHIPs that I use in essay 2, the UHS has fewer variables on educational outcomes, but it still contains sufficiently basic information on individuals' demographics, educational attainment, as well as more detailed information regarding job market outcomes (including work status, and various kinds of annual income), household wealth, and household consumption on a comprehensive list of goods.

I have access to eight rounds of the survey covering years of 2002-2009.[29] This period of time has the advantage of capturing the labor market dynamics when a surge of college-educated labor kept entering the labor market. The survey samples individuals reside in 9 provinces (representative of individuals from a total of 32 provincial districts in mainland China), except for the survey round of 2003, in which it sampled an additional 7 provinces.[30] To keep it consistent across survey years, I dropped individuals from these additional provinces for the

[29] The restricted-use UHS datasets are only open to a small groups of research teams and researchers who apply to the NBS. Professor Ning Fang at the Chinese Academy of Social Science provided me with access to these datasets.

[30] The nine provinces are: Beijing, Shanxi, Liaoning, Jiangsu, Anhui, Guangdong, Chongqing, Sichuan, and Gansu.

UHS 2003 round. For brevity, I treat the eight rounds of sample as independent repeated cross-sections, and individuals in each round as independent and identically distributed (i.i.d.).[31] On top of this dataset, I also use data from the annual Educational Statistics Yearbook of China for the years of 1971-2009 to construct macro-level measures for some education and labor market variables.[32]

3.4.2 Sample

My full analysis sample restricts to individuals born during 1963-1987 (corresponding to the 1981-2003 probable college cohorts). People older than these cohorts were affected by the Cultural Revolution during which all the higher education institutions were shut down, and people younger than these cohorts might not have completed 4-year college and worked for at least an entire year by 2009.[33] I kept individuals whose highest educational attainment is either high school (including academic and vocational high schools) or college (including 3- and 4-year colleges), and thus drop those who either have never attended upper secondary education or have attended graduate school by the time of the survey. I finally exclude those who are disabled, retired, or enrolled in school at the time of the survey. The full analysis sample pooled across eight survey rounds contain 137,476 individuals, and the sample size for each survey rounds range from 10,558 to 22,814.

---

[31] The UHS samples one thirds of the same households from the previous year, and replace all the households every three years. Therefore, it is possible that the same individuals appear more than once in my pooled sample. Moreover, individuals could be nested under households (i.e., not i.i.d.), however, I show later that after restricting the sample to certain cohorts, in most cases only one individual from each household appears in my analysis sample.

[32] Most national level annual data are available at http://data.stats.gov.cn/index.htm. Some statistics and information I use are only available in the yearbooks' hard copies, and I have inputted the data manually.

[33] Cohorts born between 1959-1962 were affected by the "3-year natural disaster" that could have distinct and long-lasting (negative) cohort effects for these individuals.

The “workers” sample further restricts to individuals who reported to be employed and earned a non-zero annual job income at the time of the survey. It thus excludes those who reported to be unemployed, worked in home, and those who were “working” but reported to have earned zero annual job income. The “workers” sample pooled across eight survey rounds contain 121,859 individuals (88.6 percent of the full analysis sample), and the sample size for each survey rounds range from 9,071 to 20,973.

3.4.3 Key Variables

Indicators for age groups. I created a set of four indicator variables to indicate whether an individual is in the age ranges of [24,27], [28,31], [32,35] or [36,39] at the time of the survey. The construction of the age groups ensures that the youngest cohort surveyed at the latest round (i.e., 2009) and the oldest cohort surveyed at the earliest round (i.e., 2002) are able to cover all the cohorts for my full analysis sample (i.e., the 1963-1987 birth cohorts, or the 1981-2003 probable college cohorts). Moreover, by starting from individuals aged 24 years old, I can ensure that graduates from 4-year colleges have worked for at least an entire year at the time of the survey. Lastly, since I later intent to construct the “age group specific supplies” variable to quantify cohort-specific supplies accrue to an individual’s own cohort as a key explanatory variable for the changes in age group specific college premiums, the range of the age groups should not be too wide (the age group specific supplies are weighted averages of cohort-specific supplies across cohorts in a given age group), nor too narrow (which would lead to large measurement error since individuals’ high school and college cohorts are inferred from their birth year and month). Nevertheless, I have experimented using different starting ages, and age gaps of 5, 7 and 10 years. The pattern of major results that I will show later is generally similar.

BA-HS wage gap. One feature of the UHS dataset is that its sole variable on "educational attainment" asks individuals to report their highest level of schooling attended, no matter one completed/obtained a degree or not. I created four dummy variables "BA" "AA" "College" and "HS" to indicate whether an individual has ever attended a regular 4-year, 3-year, either types of colleges, and high school (including both academic and vocational high schools) as their highest education level respectively. These four variables could capture individuals who at least received some education at the respective levels.

As for earnings, the UHS records four sources of annual income: employment, business, asset, and transfers. I present results using annual income from employment (only) in the main text.[34] I discount all the income variables to the 2002 Chinese currency value using annual nationwide CPIs in urban areas (NBS, 2020)

I first focus on the BA-HS wage gap for reasons explained in section 3. To compute the age group specific wage gaps for each survey year (r_{jt}), I regress log annual job income on the "BA" indicator, controlling for gender, province, and age (within each age group) fixed effects, within each age group-survey year cell for individuals whose educational attainment is either high school or 4-year college.[35] By using log income, I'm restricting my analysis to the "workers" sample. I also conduct the same sets of regressions including those with zero job earnings by adding 1 yuan to annual job income for all individuals in the full analysis sample and also including an indicator for imputed earnings.

[34] Other categories of income have higher proportion of missing values, and results are approximately identical if I instead use total income that adds up all four sources of income (assigning zeros to missing values).

[35] Alternatively, the BA-HS wage gaps could be computed as the difference in log mean annual job income between the BA and HS groups within each age group-survey year cell. However, this method is not as accurate as the "regression" approach that controls for gender, province, and age (within age groups) fixed effects.

Age specific supplies of BA- and HS-educated labor. I would like to create a variable to measure the quantity of efficient labor supply of the cohort(s) of individuals with the same education level who entered the labor market at the same time/year. Card & Lemieux (2001) measure this variable as the actual hours supplied within each education level-age group-survey year cell, constructed using data from the CPS data. Information on hours worked is unavailable for the UHS dataset, nor from any other macro-level dataset.[36] Instead, I use the number of graduates from respective education levels for each age group-survey year cell to measure the potential quantity of age specific labor supplies. In China, labor force participation rates are rather high. For example, the average labor force participation rate is 73 percent during 2002-2009 (The World Bank, 2020b). Moreover, I do not have information on the year when an individual first worked, instead, I use the year of graduation assuming that individuals would directly enter the labor market when they graduate. More specifically, I assume that people graduate from high school, 3- and 4-year colleges at 18, 21, and 22 years old respectively. As an example, age specific supply of BA-educated labor within the 24-27 age group at year 2009 ($l_{BA,2427,2009}$) would be the sum of graduates from 4-year colleges during the years of 2004-2007. Similarly, age specific supply of HS-educated labor within the 36-39 age group at year 2002 ($l_{HS,3639,2002}$) would be the sum of graduates from high school during 1981-1984. Data on the number of graduates from each educational level come from the annual Educational Statistics Yearbook of China (1971-2009). I use the number of graduates, instead of the number of degrees conferred, to better correspond to the education attainment measure used in this analysis. Given the age-group specific quantities for the two types of labor, the age specific relative supply

[36] In fact, this variable is available for the 2002-2006 rounds of the UHS, but the later rounds no longer ask this question, making it impossible to construct a panel across all the eight rounds on this variable.

variable $(\frac{l_{BA,j,t}}{l_{HS,j,t}})$ is easily computed as the ratio between BA and HS within each age group-survey year cell.

Table 3.1 presents the constructed age-group specific relative supply index by year, and Figure 3.3 shows it graphically. Age-group specific relative supplies follow divergent trends especially between the oldest and youngest age groups in my sample: this measure more than doubles for the youngest age group, but drops for the oldest age group.

Table 3.1: Age-group Specific Relative Supply Index for BA- vs. HS-educated Labor by Year.

	Survey year							
Age group	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**
Age 24-27	0.108	0.126	0.136	0.142	0.156	0.175	0.205	0.229
Age 28-31	0.071	0.074	0.078	0.093	0.108	0.126	0.136	0.142
Age 32-35	0.075	0.070	0.067	0.067	0.071	0.074	0.078	0.093
Age 36-39	0.083	0.095	0.094	0.085	0.075	0.070	0.067	0.067

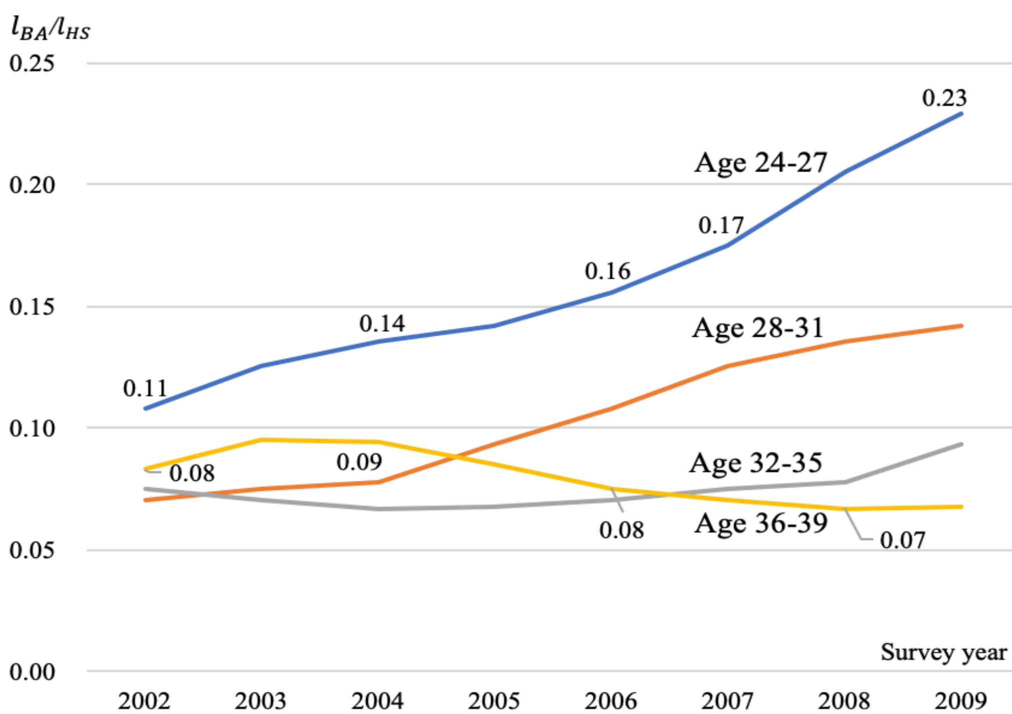


Figure 3.3: Age-group Specific Relative Supply Index for BA- vs. HS-educated Labor by Year.

Data source: Author's calculation using data from National Statistics Bureau of China (retrived from: http://www.stats.gov.cn/english/) and various years of *Educational Statistics Yearbook of China.*

Aggregate supplies of BA- and HS-educated labor. To make comparisons, I construct the aggregate supplies variables under the hypotheses of both perfect and imperfect substitution across ages. Moreover, even though my analysis of the effects of age specific supplies on wage gaps focus on the four age gaps covering individuals aged 24-39 years old, my aggregate supply measures use the sums of supplies across both the [24,39] and [16,49] age ranges.[37] The latter variable could capture the potential quantity of labor supplied from all people in legal working ages. In the following paragraphs, I explain in more details about the construction of the aggregate labor supply variables separately under the two different hypotheses.

(1) *Aggregate supplies with perfect substitution.* Given perfect substitution across age groups, the sub-aggregates of HS-educated labor ($\mathrm{L_{HS,t}}$) and BA-educated labor ($\mathrm{L_{BA,t}}$) at year t can be written as:

$$L_{HS,t} = \sum_j \left(\alpha_j l_{HS,j,t}\right), \tag{3.8}$$

and

$$L_{BA,t} = \sum_j \left(\beta_j l_{BA,j,t}\right), \tag{3.9}$$

where α_j and β_j are efficiency parameters across age groups within the HS- and BA-educated labor respectively. α_j and β_j could equal to one for all age groups, assuming that labor in different age groups supply the same efficiency unit of labor. This hypothesis could be relaxed

[37] There are two major reasons for focusing on individuals aged between 24 and 39: first, these cohorts are comparable in terms of their access to higher education institutions, and second, people older than this age range (especially those in their forties) are more likely to pursue higher education after some years of work mostly for job promotion, making the age specific supply variable (constructed from assuming people work after completing their highest level of education) inaccurate for older age groups. Reason for using the [16, 49] as alternative measure of labor supply is that the minimum legal working age in China is 16 years old, and workers can retire when they turn 50, 55, and 60 for female workers, female workers holding leadership roles, and male workers.

by adjusting the age-group specific quantities using weights (i.e., α_j and β_j) equal to the relative mean wages across age groups within the two types of labor respectively.

Besides, I measure aggregate supplies by summing up labor across both the [24,39] and [16,49] age ranges. For the former, given age specific labor supplies for the four age groups at each survey year t, the aggregate supplies are a simple (or weighted) sum of the four age groups within each education level for each survey year. The latter one sums of quantities supplied by people in all legal working ages. For example, the aggregate supply of HS-educated labor who are between the ages of 16-49 in 2002 (i.e., $L_{HS,2002}$) is the sum of high school graduates during the years of 1971-2002 (assuming high school graduates are at least 18 years old when they first enter the labor market), and the aggregate supply of BA-educated labor in the [16,49] age range in 2009 (i.e., $L_{BA,2009}$) is the sum of 4-year college graduates during the years of 1982-2009 (assuming 4-year college graduates are at least 22 years old when they first enter the labor market).

Given quantities of labor supply for the two types of labor for each year, the aggregate relative supply term in equation (3.7) (i.e., $\frac{L_{BA,t}}{L_{HS,t}}$) is easily computed as the ratio of aggregate supplies at year t between the two types of labor.

(2) Aggregate supplies with imperfect substitution. Given age group specific supplies for the two types of labor at each year, we still need to know the elasticity of substitution across age groups (i.e., σ_A) and the relative efficiency parameters across age groups for the two types of labor (i.e., $\alpha_j \; and \; \beta_j$) in order to compute aggregate supplies of BA- and HS-educated labor under imperfect substitution across ages as specified in equation (3.1) and (3.2). To estimate these parameters, I adopt the two-step estimation procedure first used by Card & Lemieux (2001). I will elaborate on this method and how I use data from the UHS to identify all the model parameters in equation (3.7) in the next sub-section.

3.4.4 Estimating the Elasticity of Substitution between Age Groups and Efficiency Parameters for Different Age Groups

Note that according to equation (3.7), the observed BA-HS wage gaps for the four age groups j= [24,27], [28,31], [32,35], and [36,39] during the sample years t= 2002, 2003, …, and 2009, are affected by a set of age-group specific factors that are constant across years [the relative efficiency parameters across age groups "ln $(\frac{\beta_j}{\alpha_j})$"], a set of year effects that are fixed across age groups [the relative technology efficiency parameters "ln $\left(\frac{\theta_{BA,t}}{\theta_{HS,t}}\right)$" and the relative aggregate supply term "$\left(\frac{1}{\sigma_A}-\frac{1}{\sigma_E}\right)$ ln $(\frac{L_{BA,t}}{L_{HS,t}})$"], and a set of age-group specific factors [the age-group specific relative supply term "$\frac{1}{\sigma_A}$ln $(\frac{l_{BA,j,t}}{l_{HS,j,t}})$"]. This implies that equation (3.7) can be re-written as:

$$r_{jt} = \gamma_j \boldsymbol{Age_j} + \delta_t \boldsymbol{Year}_t - \frac{1}{\sigma_A} \ln\left(\frac{l_{BA,j,t}}{l_{HS,j,t}}\right) + e_{jt}, \tag{3.10}$$

that is, the BA-HS wage gaps can be decomposed into age effects, survey year effects, and an additional age-group specific (or "cohort") effect. The year effects capture the combined effects of changes in both the relative supply of two types of labor, and the relative technology efficiency parameters between the two types of labor (and thus the relative demand for the two types of labor). Under perfect substitution across age groups, there will be no additional cohort effects (i.e., $\frac{1}{\sigma_A}$ = 0) after controlling for the age and year effects. We can use equation (3.10) to empirically estimate $\frac{1}{\sigma_A}$ by regressing age-group-survey-year level BA-HS wage gaps on age-group-survey-year level relative supplies of BA- vs. HS-educated labor, the four age-group dummies, and the eight survey year dummies. The coefficient on the age-specific relative supply term "ln $(\frac{l_{BA,j,t}}{l_{HS,j,t}})$" therefore gives an estimate of the elasticity of substitution across age groups (σ_A). Using the full analysis and worker samples, my estimates of $\frac{1}{\sigma_A}$ are approximately 0.04 and 0.50 respectively.

The second step is to identify the relative efficiency parameters across age groups for the two types of labor α_j and β_j given the estimate of σ_A. Note that if we take logs of both sides of equation (3.5), we can get:

$$\ln(\omega_{HS,j,t}) = \ln\varphi + \ln(\alpha_j) - \frac{1}{\sigma_A}\ln(l_{HS,j,t}), \tag{3.11}$$

where $\varphi = \theta_{HS,t}\, L_{HS,t}^{\rho-\eta}\, (\theta_{HS,t}\, L_{HS,t}^{\rho} + \theta_{BA,t}\, L_{BA,t}^{\rho})^{\frac{1-\rho}{\rho}}$ captures all the survey year effects. It could be further simplified as:

$$\ln(\omega_{HS,j,t}) + \frac{1}{\sigma_A}\ln(l_{HS,j,t}) = \epsilon_{HS,t}\boldsymbol{Year}_t + \ln(\alpha_j). \tag{3.12}$$

Similarly, re-writing equation (3.6), we can get:

$$\ln(\omega_{BA,j,t}) + \frac{1}{\sigma_A}\ln(l_{BA,j,t}) = \epsilon_{BA,t}\boldsymbol{Year}_t + \ln(\beta_j). \tag{3.13}$$

The left hand-side of equations (3.12) and (3.13) are easily computed using mean wages and supplies at each education-age group-survey year cell, and estimated from the first step. Then, we can run regressions of the sums of the two terms on the left hand-sides on the year dummies and age group dummies for HS- and BA-educated labor respectively. Coefficients on the age group dummies would give us the estimates of α_j and β_j.

Given estimates of σ_A, α_j and β_j, we can compute the aggregate supplies of the two types of labor at a given year t ($L_{HS,t}$ and $L_{BA,t}$) using equations (3.1) and (3.2), and the relative aggregate supply variable $(\frac{L_{BA,t}}{L_{HS,t}})$ in equation (3.2). Because my estimates of the $\frac{1}{\sigma_A}$ are quite large (which also aligns with the findings using data from three countries in Card & Lemieux, 2001), the aggregate relative supply indexes under perfect and imperfect substitution do not differ dramatically (I list the constructed aggregate relative supply indexes under different hypotheses by year and present these measures graphically in Table A7 and Figure A4 in Appendix A.

3.5 Decomposing the BA-HS Wage Gap

Following the literature that often assumes that there is a log-linear trend increase in the demand for skills over time caused by skill-biased technological change (Acemoglu & Autor, 2012; Goldin & Katz, 2007a; Tinbergen, 1974), the relative productivity/technology parameter between the two types of labor at year *t* can be further written as: $\ln(\frac{\theta_{ct}}{\theta_{ht}}) = \lambda_0 + \lambda_1 t$, where *t* is survey year for my case. Therefore, equation (3.7) can be further written as:

$$r_{jt} = \lambda_1 \ln\left(\frac{l_{BA,j,t}}{l_{HS,j,t}}\right) + \lambda_2 \ln\left(\frac{L_{BA,t}}{L_{HS,t}}\right) + \lambda_3 t + \lambda_4 \boldsymbol{Age_j} + \lambda_0 + e_{jt}, \tag{3.14}$$

that is, we can decompose the observed BA-HS wage gap (or the BA premium) for age group j in year t into age-group specific and aggregate relative supply factors, a linear time trend that accounts for skill-biased technological changes, and age effects (that are assumed to be fixed over time).

Table 3.2 below presents regression results using equation (3.14). The first four models use the workers sample, and the last two use the full sample. As comparisons, I also present results from using the traditional framework that assume perfect substitution across age groups in models 1 and 2. Also, the aggregate relative supply indexes are computed either using workers from the four age groups only (models 1, 3, and 5) or from all workers in legal working age (models 2, 4, and 6) under the hypothesis of either perfect or imperfect substitution across age groups.

First, comparing models that assumes imperfect substitution (models 3-6) and the traditional models that assume perfect substitution (models 1 and 2) show that models that incorporate imperfect substitution across age groups have better goodness-of-fit than models that simply sum up labor supplies of the same educational level across all the age groups.

Table 3.2: Effects of Supply, Demand, and Age on BA-HS Earnings Gaps.

	Workers Only				Full Sample	
	(1)	(2)	(3)	(4)	(5)	(6)
Age-group specific relative supply	---	---	-0.035	-0.042	-0.457	-0.531
	---	---	(0.001)	(0.001)	(0.005)	(0.005)
Aggregate relative supply with perfect	-0.430	-0.327	---	---	---	---
substitution across age groups	(0.007)	(0.020)	---	---	---	---
Aggregate relative supply with imperfect	---	---	-0.365	-0.122	-3.383	-2.459
substitution across age groups	---	---	(0.007)	(0.013)	(0.034)	0.005
Annual Temporal Trend	0.026	0.025	0.019	0.006	0.134	0.145
	(0.001)	(0.002)	(0.000)	(0.001)	(0.002)	(0.004)
Age Effects (Age 24-27 as base):						
Age 28-31	0.250	0.249	0.235	0.231	0.658	0.623
	(0.001)	(0.001)	(0.001)	(0.001)	(0.003)	(0.003)
Age 32-35	0.251	0.250	0.224	0.219	0.562	0.503
	(0.001)	(0.001)	(0.001)	(0.001)	(0.004)	(0.005)
Age 36-39	0.286	0.286	0.260	0.255	0.478	0.427
	(0.000)	(0.001)	(0.001)	(0.001)	(0.004)	(0.004)
Aggr. relative supply using age group 24-39	✔		✔		✔	
Aggr. relative supply using age group 16-49		✔		✔		✔
N	*N = 46,354*				*N = 53,750*	
R2	0.7890	0.7806	0.8902	0.8840	0.8133	0.7917

Data source: Urban Household Survey, 2002-2009.

The primary interest is the coefficient on the age-group specific relative supply variable that is estimated to be approximately -0.04 if using workers with positive earnings only (models 3 and 4). Estimates for this variable are quite robust across different models, and are also very close to the estimates of $-\frac{1}{\sigma_A}$ from the first-step of my estimation procedure. The magnitude of the coefficient means that when the supply of BA-educated labor (relative to HS-educated labor) increases by 1 percent within one's own age group, there will be an additional 0.04 percentage points decrease in the BA-HS earnings gap (i.e., BA premium) particularly for this age group, on top of the other factors that affect all age groups. Note that, age-group specific relative supply increases from 0.108 to 0.229 (increased by 112 percent) during 2002-2009 for the youngest age group (ages 24-27) in our sample, which implies that the BA premium for workers in their mid-twenties is about 4.5 percentage points lower in 2009 as compared with that for workers of the same ages in 2002, assuming that workers in both years have a job and earned positive income.

If we instead taking zero earnings into consideration, models 5 and 6 in Table 3.2 suggests that the negative cohort effects due to the increase in age-group specific relative supply are much larger: the magnitudes of the coefficients (approximately -0.5) imply that the BA premium for workers in their mid-twenties is about 56 percentage points lower in 2009 as compared with that for workers of the same age in 2002 if we also include zero earnings. Note that the validity of the coefficients for the full sample is based on the assumption that zero earnings recorded in the UHS dataset reflects true unemployment or employment without pay, not mis-reporting behavior or missing-ness, which I am unable to verify. Nevertheless, the difference in the coefficients on the age-group specific supply index between the workers and full sample might indicate that a major part of the "cohort crowding" effects work through the channel of higher probability of unemployment or working under zero earnings (e.g., some internship positions with no pays).

Figure A5 in Appendix A shows the cohort-by-cohort contrast in the BA-HS gaps in the probability of being employed and earning positive job income using data pooled across the eight rounds. Results show that the employment premiums for BA-educated labor have been above 10 percentage points (i.e., BA-educated labor has a probability of being employed and earning positive job income 10 percentage points higher than HS-educated labor), but start to drop starting from the cohort who attended college in 1996. Starting from the 2001 college cohort (i.e., two years into the college expansion), there are virtually no difference between HS- and BA-educated labor in the probability of being employed and earning a positive income.

The annual growth rate of the relative technology/efficiency parameter is consistently estimated to be around 2-3 percent for the worker sample (except for model 4, which has a lower estimate). This means college premiums increase by about 2-3 percentage points annually during 2002 to 2009 due to skill-biased technology changes. However, this temporal trend in favor of BA-educated labor is mitigated by increases in the aggregate relative supplies of BA-educated labor. Since the aggregate relative supplies grow by about 6-11 percent annually during 2002-2009, depending on different measures of the variable. Therefore, the coefficients on the aggregate relative supply indexes, coupled with corresponding changes in the respective aggregate relative supply measures, imply an annual negative effect (averaged across the 8 years) due to increases in the aggregate relative supply of BA-educated labor on the BA premiums to be around 1.5-4.3 percentage points for the worker sample, and 24.2-30.1 percent for the full sample. Combining this with the annual time trend suggests the net effect due to changes in the aggregate relative supply and demand factors on the college premiums to be about -1.8 to -0.9 percentage points for the worker sample, and -10.8 to -5.3 percentage points for the full sample, both averaged across the eight years. In other words, the BA-HS earnings gap has been narrowed down by 1-2 and 5-11 percentage points per year during 2002 to 2009 for the worker and full samples respectively, on top of any additional effects for particular age groups. Note that this downward pressure on earnings for BA-educated labor affects labor in all ages, even those whose own cohort-specific supplies do not increase. Therefore, it could be interpreted as the general equilibrium/spill-over effects from the surges in supplies of BA-educated labor from other cohorts.

When taking all the effects into consideration and conditional on having earned a positive income, for example, the combined effects of aggregate relative supply and demand factors on the BA premiums are about -9 to -6 percentage points for workers in all age groups from 2002 to 2009; the BA premiums for younger workers are further dampened by the additional cohort effects due to increases in their cohort-specific supplies. On net, BA-educated labor ages 24-27 years old enjoys a BA premium approximately 11-13 percentage points lower in 2009 than that in 2002 for BA-educated labor of the same ages. This negative net effect indicates that the shifts of the cohort-specific and aggregate relative supply curves (largely due to the college expansion) have some distributional effects: HS-educated labor is better off, especially for the younger cohorts, narrowing down the BA-HS wage gap that would be much larger if only the demand shock exists.

Moreover, we can use the fully-identified model to simulate results of alternative policies, especially the model suggests that the net effect from all changes could be positive. For example, under a special case when we assume that the relative supplies have been kept constant since 1998 (i.e., the pre-expansion level), then changes in both aggregate and age-group specific supplies are approximately 0 (i.e., assuming the educational attainment moves exactly together across cohorts), the simulated BA-HS wage gap at year 2009 could be 16-21 percentage points higher than its 1998 level for people who earn positive income.[38]

---

[38] Note that this simulation assumes that the relative technology efficiency parameter is fixed (i.e., estimated from empirical data), however, this parameter might respond endogenously to the cohort-specific and aggregate relative supplies of BA-educated labor.

Lastly, there are two cautions in interpreting the results. First, the coefficients on the age-group specific relative supply index could be biases if there are other unobserved age-group specific variables that are correlated to both the supply index and the college premium. One major potential threat is the relative ability composition of BA vs. HS labor. For example, if BA-educated labor becomes less able (compare with HS-educated labor) as relative supply of BA-educated labor increases, and if we assume that relative ability is positively correlated to the college premium, then the coefficient on the age-group specific supply variable will be downwardly biased. Intuitively, as larger shares of students start to attend college, the average ability of college-educated labor could decline. However, the average ability level of HS-educated labor would drop as well since more able students shift from high school graduates only to achieve higher educational attainment. Consequently, it is not obvious how the relative ability of BA- vs. HS-educated labor would change as the relative labor supply of BA-educated labor increases. I explore this issue using two methods. I first regress the age-group-survey-year level BA-HS gaps on dummies for individuals' college cohorts, controlling for age-group specific relative supply, age effects, and survey year effects. The cohort effects are statistically significant for less than half of the cohorts, though the magnitudes of the estimates are relatively small (nearly all are less than 0.01.) Also, there is no clear pattern of the cohort effects, and particularly for the after-expansion cohorts that would be most likely to suffer from declines of either the quality of college education or average ability level of peers (or both), the cohort effects are in fact almost all positive (using the oldest college cohort of 1981 as the base). Lastly, the coefficient on the relative labor supply variable (i.e., 0.031) becomes slightly lower (not higher), and is still close to the estimates presented in Table 3.2, after controlling for additional cohort effects.

I also experiment to use proxies to measure the ability composition of BA vs. HS labor. I use cohort-level ratios between HS/BA graduates and the total population within a certain cohort to measure cohort-specific average ability levels of BA- and HS-educated labor respectively, and then use the ratio between the two ability levels to proxy for their relative ability. The coefficients on this relative ability measure are not statistically significant at the 0.05 level, and the estimates for other model parameters seldom change. Results from both methods are reassuring in that they suggest the observed effects of age-group specific relative supply on the college premiums are not driven by changes in the relative ability composition of BA vs. HS-educated labor across years and age groups.

Second, my analysis is based on the hypothesis that there is one labor market for the entire China, i.e., there are two unified equilibrium wage rates for HS- and BA-educated labor in all the provinces. Even this is not the case, the small proportion of individuals who have ever migrated across provinces in the UHS dataset (2.45 and 2.81 percent for the workers and full samples), and my controls for province fixed effects could help eliminate the major influences of arbitrage from selective migration. Nevertheless, I replicate Table 3.2 excluding individuals who reported to have ever changed their "hukou" province by the time of the survey. Results hardly change.

3.6 Discussion and Conclusion

This study uses multiple repeated cross-sections of data to explore how, at the macro-level, has the 4-year-college-high school (BA-HS) wage gap evolved after China's college expansion. This paper is motivated by the observation that the BA-HS wage gaps do not rise and fall in a parallel manner for workers of different age groups during the 2000s. Data from a nationally representative survey - the Urban Household Survey - for the years between 2002 and 2009 show that on average, the gap in logarithm of annual job earning between BA- and HS-educated workers declines from 61.1 percent in 2002 to 57.9 percent in 2009 (a 5.0 percent decrease). However, this gap was slightly widened for older workers (a 10 percent increase for workers aged 40-43 years old, and 8 percent for those aged 32-39), but shrink substantially for younger workers (a 23 percent decrease for those aged 24-27, and 10 percent for those aged 28-31). Incorporating a model with imperfect substitution among workers with the same educational level but in different age groups, I show that this differential pattern and the associated decreasing returns to 4-year college for younger cohorts correspond to surges in the cohort-specific relative supply of college-educated workers stimulated by China's college expansion. More specifically, results from decomposing the age-group-survey-year level BA-HS wage gaps into age, year, and cohort effects suggest that a 1 percent increase in the relative supply of BA-educated workers within one's own cohorts would depress the BA-HS wage gap by 0.04 percentage point. Given that college enrollment increased by nearly 4 times from 1998 to 2005, the negative cohort effects could be substantial: for example, the cohort-specific relative supply for the youngest age group in my analysis increased by 112 percent from 2002 to 2009, suggesting an additional 4.5 percentage points decrease in the BA-HS wage gap for workers of this particular age group, on top of the effects of changes in aggregate relative supply and demand that are borne by workers in all age groups. Moreover, my estimates suggest a steadily increasing relative demand for BA-educated labor at an annual growth rate of around 2-3 percentage points; it is mitigated by the negative effects from the increase in the aggregate relative supply of BA-educated labor though; the latter effect also implies that the expansion has negative spillover effects on workers who attended college before the expansion.

My findings have some implications: first, the cohort-specific relative supply of BA-educated labor due to shift in the average educational attainment for certain cohort(s) is the major explanation of the variations in the observed BA premiums by age groups and across years. this kind of large-scale variation in education attainment that affect the entire cohort(s) of individuals is essentially different from a micro intervention that make one single or a small number of individuals to obtain more education, holding all other conditions constant. The substantial changes in the cohort-specific relative supply of BA-educated labor also modify the aggregate relative supply of BA-educated labor, leading to general-equilibrium effect that affects all the workers.

Second, the growth of the relative demand for BA-educated labor is another key factor for determining college premiums. Note that the average annual GDP growth rate is around 10.8 percent during the period under this study, and without this rapid economic growth that mitigate the negative effects from increases in the relative supply of BA-educated labor, the BA-HS gap could be further depressed.

Third, the commonly-used Mincerian returns estimated using cross-sectional data could be very different from cohort-based estimates of the college premiums. Note that the former one is essentially the (weighted) average of the premiums across all cohorts at one time, and when there is large variation in the premiums among different cohorts, the Mincerian returns might not be a good indicator to guide individual decisions about human capital investment, nor to inform any cohort-specific public policies.

Lastly, when more recent rounds of earnings data become available, it is interesting to study how does the influence of cohort-specific relative supply on the college premiums change as workers grow older. Particularly, it is worthwhile to investigate whether the negative cohort effects on the youngest age group found in this analysis fade away, persist, or even amplified as these “unlucky cohorts” age and gain more years of working experience. Empirical studies conducted in the U.S. suggest that the negative effects for cohort(s) who graduate and enter the labor market at a regression or worse economy could be long-lasting and large, persist for ten years (Oreopoulos, Wachter, & Heisz, 2012) or even two decades (Kahn, 2010).

References

Abadie, A., Diamond, A., & Hainmueller, J. (2015). Comparative politics and the synthetic control method. *American Journal of Political Science 59* (2): 495–510.

Acemoglu, D. (1998). Why do new technologies complement skills? Directed technical change and wage inequality. *The Quarterly Journal of Economics*, *113*(4), 1055-1089.

———. (2010). Theory, general equilibrium, and political economy in development economics. *Journal of Economic Perspectives*, *24*(3), 17-32.

Acemoglu, D., & Angrist, J. D. (2001). How large are human-capital externalities? Evidence from compulsory-schooling laws. In B. Bernanke, & K. Rogoff (Eds.), *NBER Macroeconomics Annual*, 9–59. Cambridge: National Bureau of Economic Research.

Acemoglu, D., & Autor, D. (2010). Skills, tasks and technologies: Implications for employment and earnings. In D. Card & O. Ashenfelter (Eds.), *Handbook of Labor Economics, 4B,* 1043-1171. Elserier: San Diego and Amsterdam, North-Holland.

Acemoglu, D., Autor, D. H., & Lyle, D. (2004). Women, war, and wages: The effect of female labor supply on the wage structure at midcentury. *Journal of political Economy*, *112*(3), 497-551.

Acemoglu, D., Johnson, S., Querubin, P. & Robinson, J.A. (2008). *When does policy reform work? The case of central bank independence* (No. w14033). National Bureau of Economic Research.

Acemoglu, D., & Restrepo, P. (2018). The race between machine and man: Implications of technology for growth, factor shares and employment. *American Economic Review*, *108*(6), 1488-1542.

Aghion, P., Boustan, L., Hoxby, C., & Vandenbussche, J. (2009). The causal impact of education on economic growth: evidence from US. *Brookings papers on economic activity*, *1*, 1-73.

Altonji, J. G., & Zimmerman, S. D. (2017). The costs of and net returns to college major. In *Productivity in Higher Education*. University of Chicago Press.

Angrist, J.D. (1995). The economic returns to schooling in the West Bank and Gaza Strip. *The American Economic Review*, 1065-1087.

Angrist, J. D., & Imbens, G. W. (1995). Two-stage least squares estimation of average causal effects in models with variable treatment intensity. *Journal of the American statistical Association*, *90*(430), 431-442.

Angrist, J. D., & Krueger, A. B. (2001). Instrumental variables and the search for identification: From supply and demand to natural experiments. *Journal of Economic perspectives, 15*(4): 69-85.

Angrist, J. D., & Pischke, J. S. (2014). *Mastering Metrics: The Path from Cause to Effect.* Princeton University Press.

Appleton, S., Song, L. & Xia, Q. (2014). Understanding urban wage inequality in China 1988–2008: Evidence from quantile analysis. *World Development*, *62*, 1-13.

Autor, D. H. (2014). Skills, education, and the rise of earnings inequality among the "other 99 percent." *Science, 344*(6186), 843–851.

Autor, D. H., Katz, L. F., & Krueger, A. B. (1998). Computing inequality: Have computers changed the labor market." *The Quarterly Journal of Economics*, *113*(4), 1169-1213.

Autor, D. H., Katz, L. F., & Kearney, M.S. (2008). Trends in US wage inequality: Revising the revisionists. *The Review of economics and statistics*, *90*(2), 300-323.

Autor, D. H., Manning, A., & Smith., C. L. (2010). *The Contribution of the Minimum Wage to U.S. Wage Inequality over Three Decades: A Reassessment*, NBER Working Paper No. 16533.

Autor, D. H, Levy, F., & Murnane, R. (2003). The skill content of recent technological change: An empirical exploration. *The Quarterly Journal of Economics*, *118*(4), 1279-1333.

———. (2015). The contribution of the minimum wage to U.S. wage inequality over three decades: A reassessment. *American Economic Journal: Applied Economics*, *8*(1), 58-99.

Athey, S., & Imbens, G. W. (2017). The state of applied econometrics: Causality and policy evaluation. *Journal of Economic Perspectives 31* (2), 3–32.

Barr, A., & Turner, S.E. (2013). Expanding enrollments and contracting state budgets: The effect of the Great Recession on higher education. *The ANNALS of the American Academy of Political and Social Science*, *650*(1), 168-193.

Beaudry, P., Green, D. A., & Benjamin, M. S. (2014). "The declining fortunes of the young since 2000. *American Economic Review Papers & Proceedings*, *104*(5), 381-386.

Becker, G. S. (1962). Investment in human capital: A theoretical analysis. *Journal of political economy*, *70*(5, Part 2), 9-49.

———. (1994). Human capital revisited. In *Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education (3rd Edition)*, 15-28. The university of Chicago press.

Becker, G. S., & Chiswick, B. R. (1966). Education and the Distribution of Earnings. *The American Economic Review*, *56*(1/2), 358-369.

Berman, E., Bound, J., & Griliches, Z. (1994). Changes in the demand for skilled labor within US manufacturing: evidence from the annual survey of manufactures. *The Quarterly Journal of Economics, 109*(2), 367-397.

Bianchi, N. (2015). *The general equilibrium effects of educational expansion*. SIEPR Discussion Paper No. 15-008. Stanford, CA: Stanford Institute for Economic Policy Research.

Black, D.A., & Smith, J.A. (2004). How robust is the evidence on the effects of college quality? Evidence from matching. *Journal of Econometrics, 121*(1), 99-124.

Blau, F., & Kahn, L. M. (2005). Do cognitive test scores explain higher U.S. wage inequality? *The Review of Economics and Statistics, 87*(1), 184-193.

Bound, J., Kzedi, G., & Turner, S. (2000). Trade in university training: cross state variation in the production and use of college educated labor. MIMEO, University of Michigan.

Bound, J., & Turner, S. (2007). Cohort crowding: How resources affect collegiate attainment. *Journal of public Economics*, *91*(5-6), 877-899.

Brand, J.E., & Halaby, C. N. (2006). Regression and matching estimates of the effects of elite college attendance on educational and career achievement. *Social Science Research, 35*(3), 749-770.

Brand, J. E., & Xie, Y. (2010). Who benefits most from college? Evidence for negative selection in heterogeneous economic returns to higher education. *American sociological review*, *75*(2), 273-302.

Branstetter, L. & Lardy, N. (2008). China's embrace of globalization. In L. Brandt, & T. G. Rawski, (Eds.), *China's Great Economic Transformation*, 633-83, Cambridge: Cambridge University Press.

Brinch, C. N., Mogstad, M., & Wiswall, M. (2017). Beyond LATE with a discrete instrument. *Journal of Political Economic, 125*(4): 985–1039.

Brown, C. (1999). Minimum wages, employment, and the distribution of income. In O. Ashenfelter & D. Card (Eds.), *Handbook of Labor Economics, 3B*, 2101- 2163. Elsevier.

Byron, R. P., & Manaloto, E. Q. (1990). Returns to education in china. *Economic Development and Cultural Change 38*(4), 783–796.

Card, D. (1995). Earnings, schooling, and ability revisited. in S. Polachek (Eds.), *Research in Labor Economics*, *14*, Greenwich CT: JAI Press.

Card, D. (1999). The causal effect of education on earnings. In *Handbook of Labor Economic*s, 3, 1801-1863, Elsevier.

Card, D. (2001). The effects of unions on wage inequality in the US labor market. *Industrial and Labor Relations Review*, *54*(2), 296-315.

Card, D., & DiNardo, J. (2002). Skill-biased technological change and rising wage inequality: Some problems and puzzles. *Journal of Labor Economics*, *20*(4), 733-783.

Card, D., & Lemieux, T. (2001). Can falling supply explain the rising return to college for younger men? A cohort-based analysis. *The Quarterly Journal of Economics*, *116*(2), 705-746.

Carneiro, P., & Lee, S. (2011). Trends in quality-adjusted skill premia in the United States, 1960-2000. *American Economic Review 101*(6), 2309-49.

Carnoy, M., Loyalka, P.K., Androushchak, G.V. & Proudnikova, A. (2012). The economic returns to higher education in the BRIC countries and their implications for higher education expansion. *Higher School of Economics Research Paper,* No. WP BRP, 2.

Chakrabarti, R., Gorton, N., & Lovenheim, M.F. (2017). The effect of state funding for postsecondary education on long-run student outcomes. *IZA Conference paper.*

Che, Y., & Zhang, L. (2018). Human capital, technology adoption and firm performance: Impacts of China's higher education expansion in the late 1990s. *The Economic Journal*, *128*(614), 2282-2320.

Currie, J., & Moretti, E. (2003). Mother's education and the intergenerational transmission of human capital: Evidence from college openings. *The Quarterly journal of economics*, *118*(4), 1495-1532.

Dale, S.B., & Krueger, A.B. (2002). Estimating the payoff to attending a more selective college: An application of selection on observables and unobservables. *The Quarterly Journal of Economics*, *117*(4), 1491-1527.

———. (2011). Estimating the return to college selectivity over the career using administrative earnings data. Working paper. National Bureau of Economic Research.

Dee, T. S. (2004). Are there civic returns to education?. *Journal of Public Economics*, *88*(9-10), 1697-1720.

DiNardo, J., & Lee, D. S. (2011). Program evaluation and research designs. In A. Ashenfelter & D. Card (Eds.), *Handbook of Labor Economics, 4A*, chapter 5, 463–536. Elsevier Science B.V.

DiNardo, J., Fortin, N. M., & Lemieux, T. (1996). Labor market institutions and the distribution of wages,1973-1993: A semi-parametric approach. *Econometrica*, *64*(5), 1001-1044.

Duflo, E. (2001). Schooling and labor market consequences of school construction in Indonesia: Evidence from an unusual policy experiment. *American Economic Review 91*(4), 795-813.

———. (2004). The medium run effects of educational expansion: Evidence from a large school construction program in Indonesia. *Journal of Development Economics 74*(1), 163-197.

Duflo, E., Dupas, P., & Kremer, M. (2011). Peer effects, teacher incentives, and the impact of tracking: Evidence from a randomized evaluation in Kenya. *American Economic Review*, *101*(5), 1739-74.

Edlund, L., & Kopczuk, W. (2009). Women, wealth, and mobility. *American Economic Review*, *99*(1), 146-78.

Educational Statistics Yearbook of China. various years: 1971-2016. People's Education Press.

Fan, E., Meng, X., Wei, Z., & Zhao. G. (2010). Rates of return to university education: The regression discontinuity design. *IZA Discussion Paper No. 4749.*

Fernandez, R., & Rogerson, R. (1998). Public education and income distribution: A dynamic quantitative evaluation of education-finance reform. *American Economic Review 88*(4), 813-833.

Figlio, D. N., & Page, M. E. (2002). School choice and the distributional effects of ability tracking: does separation increase inequality?. *Journal of Urban Economics*, *51*(3), 497-514.

Fishlow, A. (1966). The American Common School Revival: Fact or Fancy? In H. Rosovsky (Eds.), *Industrialization in Two Systems: Essays in Honor of Alexander Gerschenkron.* New York: Wiley.

Gao, W., & Smyth, R. (2015). Education expansion and returns to schooling in urban china, 2001–2010: Evidence from three waves of the china urban labor survey. *Journal of the Asia Pacific Economy 20* (2), 178–201.

Ge, S., & Yang, D. T. (2011). Labor market developments in China: A neoclassical view. *China Economic Review 22*(4), 611–625.

Giles, J., Park, A., & Wang, M. (2015). The great proletarian cultural revolution, disruptions to education, and the returns to schooling in urban china. The World Bank.

Goldin, C., & Katz, L.F. (2007a). “Long-run changes in the wage structure: Narrowing, widening, polarizing.” *Brookings Papers on Economic Activity, 2,* 135-165.

———. (2007b). *The race between education and technology: the evolution of US educational wage differentials, 1890 to 2005* (No. w12984). National Bureau of Economic Research.

———. (2009). *The Race between Education and Technology.* Harvard University Press.

Gonzlez, N., & Oyelere, R. U. (2011). Are returns to education on the decline in Venezuela and does Mission Sucre have a role to play? *Economics of Education Review, 30*, 1348-1369.

Griliches, Z. (1977). Estimating the returns to schooling: Some econometric problems. *Econometrica: Journal of the Econometric Society*, 1-22.

Gustafsson, B., Shi, L., & Sato, H. (2014). Data for studying earnings, the distribution of household income and poverty in China. *China Economic Review*, *30*, 419-431.

Haider, S. & Solon, G. (2006). Life-cycle variation in the association between current and lifetime earnings. *American Economic Review, 96*(4), 1308-1320.

Heckman, J., Humphries, J. E., & Veramendi, G. (2018). Returns to education: The causal effects of education on earnings, health, and smoking. *Journal of Political Economy*, *126*(S1), S197-S246.

Heckman, J., Lochner, L. & Taber, C. (1998a). General equilibrium treatment effects: A study of tuition policy. *American Economic Review, Papers and Proceedings, 88*(2), 381-386.

———. (1998b). Tax policy and human capital formation. *American Economic Review, Papers and Proceedings, 88*(2), 293-297.

Heckman, J., Lochner, L., & Todd, P.E. (2003). *Fifty years of Mincer earnings regressions* (No. w9732). National Bureau of Economic Research.

———. (2006). Earnings functions, rates of return and treatment effects: The Mincer equation and beyond. In E. Hanushek, & F. Welch (Eds.), *Handbook of the Economics of Education,* 307–458. Rotterdam: Elsevier.

———. (2008). Earnings functions and rates of return. *Journal of Human Capital, 2*(1), 1–31.

Heckman, J., & Robb, R. (1985). Alternative methods for estimating the impact of interventions. In J. Heckman, & b. Singer (Eds.), *Longitudinal Analysis of Labour Market Data.* Cambridge: Cambridge University Press.

Heckman, J., & Vytlacil, E. (2005). Structural equations, treatment effects and econometric policy evaluation. *Econometrica, 73*(3), 669–738.

Hoekstra, M. (2009). The effect of attending the flagship state university on earnings: A discontinuity-based approach. *The Review of Economics and Statistics*, *91*(4), 717-724.

Hoxby, C. M. (2017). The productivity of us postsecondary institutions. In *Productivity in Higher Education*. University of Chicago Press.

Hu, A., & Hibel, J. (2015). Increasing heterogeneity in the economic returns to higher education in urban China. *The social science journal*, *52*(3), 322-330.

Hu, C. 2018. Three essays on college earnings premium and china's higher education expansion. Dissertation.

Huang, Z. (2001). Cong guoji gaodeng jiaoyu dazonghua kan woguo gaoxiao kuozhao zhengce (in Chinese. translated as "Understanding the expansion of higher education in China from the perspective of mass higher education internationally."). *Zhongguo Jiaoyu Zhengce Pinglun (Review of Chinese Educational Policy*).

Imbens, G. W., & Angrist, J. D. (1994). Identification and estimation of local average treatment effects. *Econometrica, 62*(2), 467-475.

Imbens, G. W., & Rubin, D. B. (2015). *Causal Inference in Statistics, Social, and Biomedical Sciences.* Cambridge University Press.

Imbens, G. W., & Wooldridge, J. M. (2009). Recent developments in the econometrics of program evaluation. *Journal of Economic Literature 47* (1): 5–86.

Jackson, C.K., Johnson, R.C., & Persico, C. (2016). The effects of school spending on educational and economic outcomes: Evidence from school finance reforms. *The Quarterly Journal of Economics, 131*(1), 157-218.

Jacob, B. A., & Lefgren, L. (2004). Remedial education and student achievement: A regression-discontinuity analysis. *Review of economics and statistics*, *86*(1), 226-244.

Jepsen, C., Troske, K., & Coomes, P. (2014). The labor-market returns to community college degrees, diplomas, and certificates. *Journal of Labor Economics*, *32*(1), 95-121.

Jia, R., & Li, H. (2016). Access to elite education, wage premium, and social mobility: The truth and illusion of China's college entrance exam. Stanford Center for International Development, Working Paper 577.

Johnson, E. N., & Chow, G. C. (1997). Rates of return to schooling in China. *Pacific Economic Review 2*(2), 101–113.

Johnson, G. E. (1997). Changes in earnings inequality: the role of demand shifts. *Journal of economic perspectives, 11*(2), 41-54.

Juhn, C., Murphy, K. M., & Pierce, B. (1993). Wage inequality and the rise in returns to skill. *The Journal of Political Economy*, *101*(3), 410-442.

Kaboski, J. J., & Townsend, R.M. (2011). A structural evaluation of a large-scale quasi-experimental microfinance initiative. *Econometrica 79*(5), 1357–1406.

Kahn, L. B. (2010). The long-term labor market consequences of graduating from college in a bad economy. *Labour economics*, *17*(2), 303-316.

Katz, L. F., & Autor, D. H. (1999). Changes in the wage structure and earnings inequality. In *Handbook of Labor Economics, 3,* 1463-1555.

Katz, L.F., & Murphy, K.M. (1992). Changes in relative wages, 1963 – 1987: supply and demand factors. *Quarterly Journal of Economics, 107*(1), 35–78.

Khanna, G. (2015). Large-scale education reform in general equilibrium: Regression discontinuity evidence from India. unpublished paper, University of Michigan.

Kline, P., & Walters, C. (2016). Evaluating public programs with close substitutes: The case of Head Start. *Quarterly Journal of Economics, 131*(4), 1795–1848.

Kopczuk, W., Saez, E., & Song, J. (2010). Earnings inequality and mobility in the United States: evidence from social security data since 1937. *The Quarterly Journal of Economics*, *125*(1), 91-128.

Kyui, N. (2016). Expansion of higher education, employment and wages: Evidence from the Russian Transition. *Labour Economics, 39*, 68-87.

Lavy, V., Paserman, M. D., & Schlosser, A. (2012). Inside the black box of ability peer effects: Evidence from variation in the proportion of low achievers in the classroom. *The Economic Journal*, *122*(559), 208-237.

Lee, D., & Wolpin, K. (2006). Intersectoral labor mobility and the growth of the service sector. *Econometrica, 74*(1), 1-46.

Lemieux, T. (2006). Increasing residual wage inequality: Composition effects, noisy data, or rising demand for skill? *The American Economic Review*, *96*(3), 461-498.

Lemieux, T., & Card, D. (2001). Education, earnings, and the 'Canadian GI Bill'. *Canadian Journal of Economics, 34*(2), 313-344.

Li, H., Liang, J., & Wu, B. (2016). Labor market experience and returns to education in fast growing Economies. Unpublished paper, Tsinghua University.

Li, H., Liu, P. W., & Zhang, J. (2012). Estimating returns to education using twins in urban China. *Journal of Development Economics, 97*(2), 494–504.

Li, L. (2003). Li Lanqing jiaoyu fangtanlu (In Chinese. translated as "Interviews on education with Li Lanqing."). Beijing: People's Education Press.

Li, S., Whalley, J., & Xing, C. (2014). China's higher education expansion and unemployment of college graduates. *China Economic Review*, *30*, 567-582.

Liu, Z. (1998). Earnings, education, and economic reforms in urban china. *Economic Development and Cultural Change, 46*(4), 697–725.

Lochner, L., & Moretti, E. (2004). The effect of education on crime: Evidence from prison inmates, arrests, and self-reports. *American economic review*, *94*(1), 155-189.

Long, M.C. (2008). College quality and early adult outcomes. *Economics of Education Review, 27*(5), 588-602.

Low, H., & Meghir, C. (2017). The use of structural models in econometrics. *Journal of Economic Perspectives, 31*(2), 33-58.

Loyalka, P. (2009). Three essays on Chinese higher education after expansion and reform: Sorting, financial aid, and college selectivity. Dissertation.

Meng, X. (2012). Labor market outcomes and reforms in china. *The Journal of Economic Perspectives, 26*(4), 75–101.

Min, W. (2004). The legacy of the past and the context of the future. In P. G. Altbach & T. Umakoshi (Eds.). *Asian Universities: Historical Perspectives and Contemporary Challenges.* Baltimore: Johns Hopkins University Press.

Minaya, V., & Scott-Clayton, J. (2016). *Labor Market Outcomes and Postsecondary Accountability: Are Imperfect Metrics Better than None?* NBER Working Paper No. w22880. National Bureau of Economic Research.

Mincer, J. (1974). *Schooling, Experience, and Earnings.* New York: Columbia University Press.

Ministry of Education of the People's Republic of China [MoE]. (1996). *Opinion on Management of Admission Plan of Universities and Colleges.* (in Chinese).

———. (2006). *National Statistical Report on Education Development.* (in Chinese).

———. (2020a). Statistics: Higher Education. http://www.moe.gov.cn/jyb_sjzl/sjzl_fztjgb/201707/t20170710_309042.

———. (2020b). Statistics: Higher Education. http://www.moe.gov.cn/s78/A03/moe_560/jytjsj_2015/2015_qg/201610/t20161012_284510.

Mirowsky, J. (2017). *Education, social status, and health*. Routledge.

Murnane, R. J., & Willett, J. B. (2011). *Methods Matter: Causal Inference in Educational and Social Science Research.* New York: Oxford University Press.

Murphy, R., Scott-Clayton, J., & Wyness, G. (2017). The end of free college in England: Implications for quality, enrolments, and equity. NBER Working Paper No. w23888. National Bureau of Economic Research.

National Bureau of Statistics of China [NBS]. (2020). National Data. retrieved from: http://data.stats.gov.cn/index.htm.

National Center for Education Statistics. (2019). Undergraduate Enrollment. retrieved from: https://nces.ed.gov/programs/coe/indicator_cha.asp.

National Student Clearinghouse Research Center. (2019). *Undergraduate Degree Earners Report 2017-18.*

Noorbakhsh, F., Paloni, A., & Youssef, A. (2001). Human capital and FDI inflows to developing countries: New empirical evidence. *World development*, *29*(9), 1593-1610.

Oreopoulos, P., & Petronijevic, U. (2013). Making college worth it: A review of the returns to higher education. *The Future of Children*, *23*(1), 41–65.

Oreopoulos, P., & Salvanes, K. G. 2011. Priceless: The nonpecuniary benefits of schooling. *Journal of Economic Perspectives, 25*(1), 159–184.

Oreopoulos, P., Von Wachter, T., & Heisz, A. (2012). The short-and long-term career effects of graduating in a recession. *American Economic Journal: Applied Economics*, *4*(1), 1-29.

Ost, B., Pan, W., & Webber, D. (2018). The returns to college persistence for marginal students: regression discontinuity evidence from university dismissal policies. *Journal of Labor Economics*, *36*(3), 779-805.

Ou, D., & Zhao, Z. (2016). Higher education expansion and labor market outcomes for young college graduates. IZA Discussion Paper No. 9643.

Pack, H., & Nelson, R.R. (1999). *The Asian miracle and modern growth theory*. The World Bank.

Peking University School of Education. (2011). *Employment Status of Higher Education Graduates in China, 2009-2010*, Beijing: Peking University Press.

Piketty, T., & Saez, E. (2003). Income inequality in the United States, 1913-1998. *The Quarterly Journal of Economics*, *118*(1), 1-39.

Ridley, M., & Terrier, C. (2018). Fiscal and education spillovers from charter school expansion. NBER Working Paper No. w25070. National Bureau of Economic Research.

Rubin, D.B. (1978). Bayesian inference for causal effects: The role of randomization. *The Annals of statistics*, 34-58.

Schlotter, M., Schwerdt, G., & Woessmann, L. (2009). Methods for causal evaluation of education policies and practices: An econometric toolbox. European Expert Network on Economics of Education (EENEE).

Schneider, B., Carnoy, M., Kilpatrick, J., Schmidt, W.H., & Shavelson, R.J. (2007). *Estimating Causal Effects using Experimental and Observational Designs: A Think Tank White Paper.* Washington, D.C.: American Educational Research Association.

Spitz-Oener, A. (2006). Technical change, job tasks and rising educational demands: Looking outside the wage structure. *Journal of Labor Economics*, *24*(2), 235-270.

Svensson, J., & Yanagizawa-Drott, D. (2012). Estimating impact in partial vs. general equilibrium: A cautionary tale from a natural experiment in Uganda. Working paper.

The World Bank. (2020a). Data: School enrollment, tertiary (% gross). retrieved from: https://data.worldbank.org/indicator/SE.TER.ENRR.

———. (2020b). Data: Labor force participation rate, total (% of total population ages 15+). retrieved from: https://data.worldbank.org/indicator/SL.TLF.CACT.ZS?locations=CN.

Tinbergen, J. (1974). Substitution of graduate by other labour. *Kyklos: International Review for Social Sciences*.

Trow, M. (1972). Technical report: Carnegie Commission national survey of higher education. Ann Arbor: Inter- university Consortium for Political and Social Research.

Valero, A., & Van Reenen, J. (2016). *The economic impact of universities: Evidence from across the globe.* (No. 1444). Center for Economic Performance Research, London School of Economics and Political Science, London, UK.

VanderWeele, T. (2015). *Explanation in Causal Inference: Methods for Mediation and Interaction.* Oxford University Press, New York.

Voena, A. (2015). Yours, mine, and ours: Do divorce laws affect the intertemporal behavior of married couples? *American Economic Review. 105*(8), 2295–2332.

Wu, Y., & Zhao, Q. (2010). Higher Education Expansion and Employment of University Graduates. *Economic Research Journal*, *9*, 93-108.

Xing, C., Yang, P., & Li, Z. (2018). The medium-run effect of China's higher education expansion on the unemployment of college graduates. *China Economic Review*, *51*, 181-193.

Yang, J., & Gao, M. (2018). The impact of education expansion on wage inequality. *Applied Economics, 50*(12), 1309-1323.

Yano, S. (2012). *Overeducated? The impact of higher education expansion in post-transition Mongolia*. New York, NY: Columbia University. Dissertation

Yu, L. (2017). Did better colleges bring better jobs? Estimating the effects of college quality on initial employment for college graduates in China. *Current Issues in Comparative Education, 19*(2): 166-197.

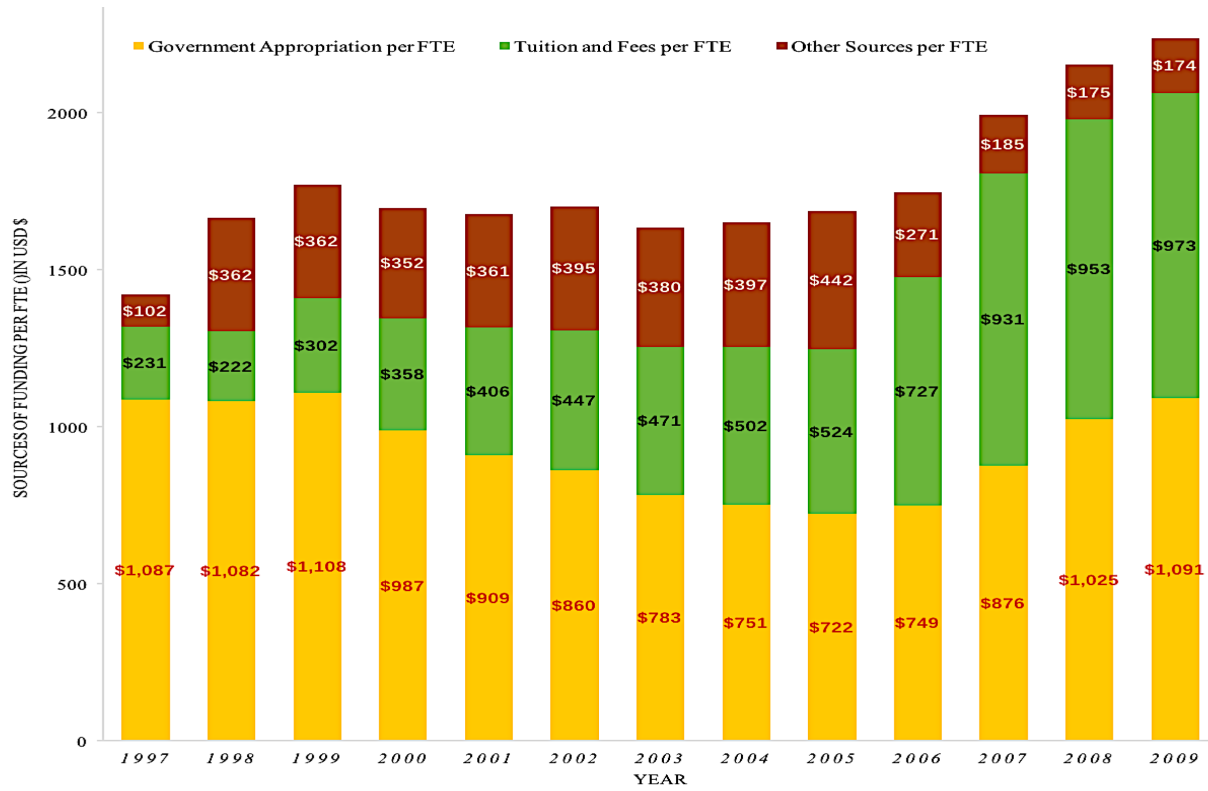
Zhang, J., Zhao, Y., Park, A., & Song, X. (2005). Economic returns to schooling in urban China, 1988 to 2001. *Journal of comparative economics, 33*(4), 730–752.

Zhou, S. (2017). Higher education in China: Inequality in access, college allocation behaviors and the effect of financial aid. Dissertation.

Zimmerman, S. D. (2014). The returns to college admission for academically marginal students. *Journal of Labor Economics*, *32*(4), 711-754.

Appendix A: Appendix Figures and Tables

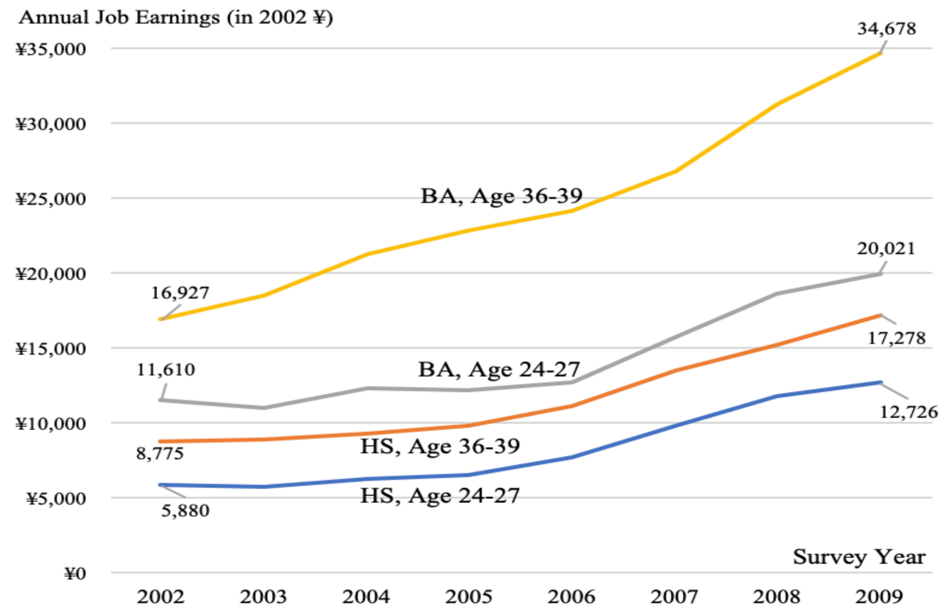
Figure A1: Sources of Funding-per-student in Regular Higher Education Institutions (Converted to USD), 1997-2009.



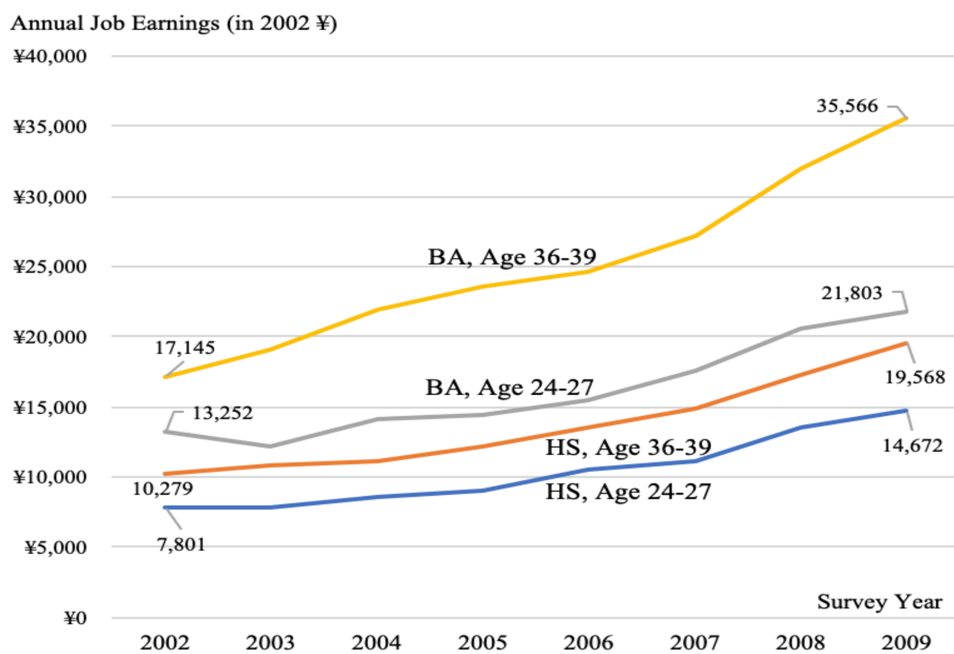
Data source: Author's calculation using data from various years of *Educational Statistics Yearbook of China.*

Figure A2: Annual Job Earnings by Selected Education Levels and Age Groups.

A. Full Sample



B. Workers Earning Positive Income

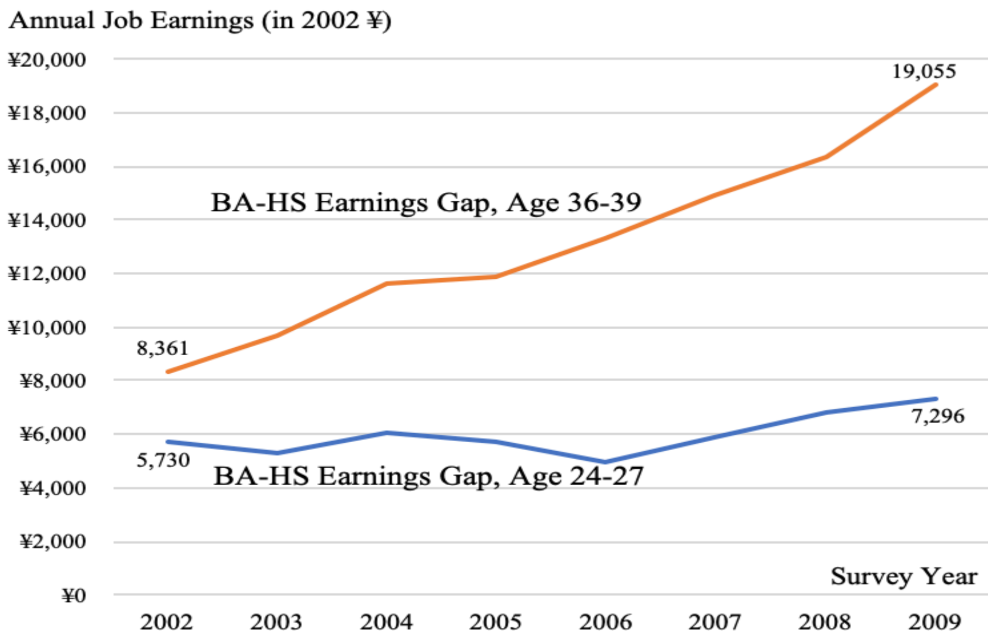


Data source: Urban Household Survey, 2002-2009.

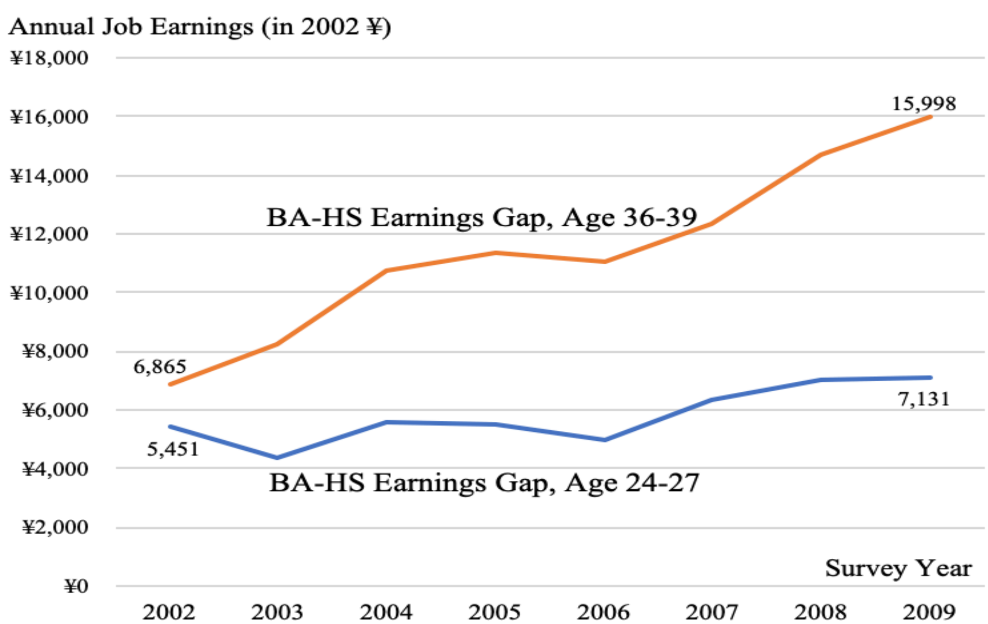
Notes: Annual incomes are discounted to the 2002 Chinese currency (¥) value using annual nationwide CPIs in urban areas obtained from the China Statistical Yearbooks, 2002-2009.

Figure A3: BA-HS Earnings Gap by Selected Age Groups.

A. Full Sample



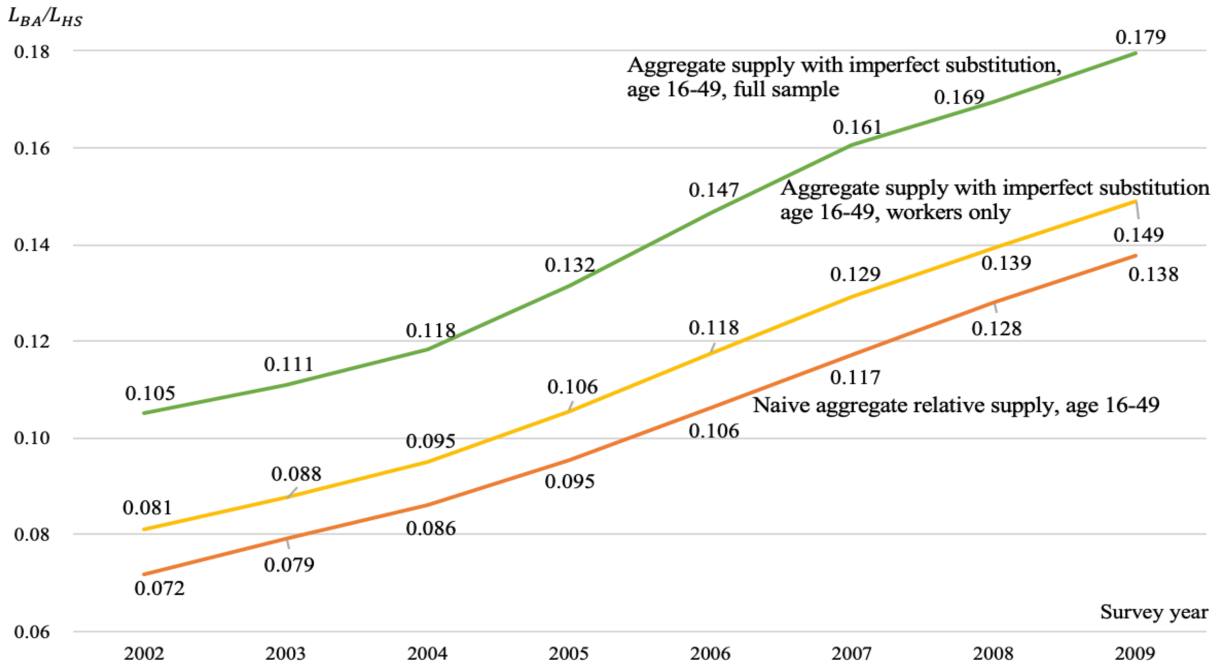
B. Workers Earning Positive Income



Data source: Urban Household Survey, 2002-2009.

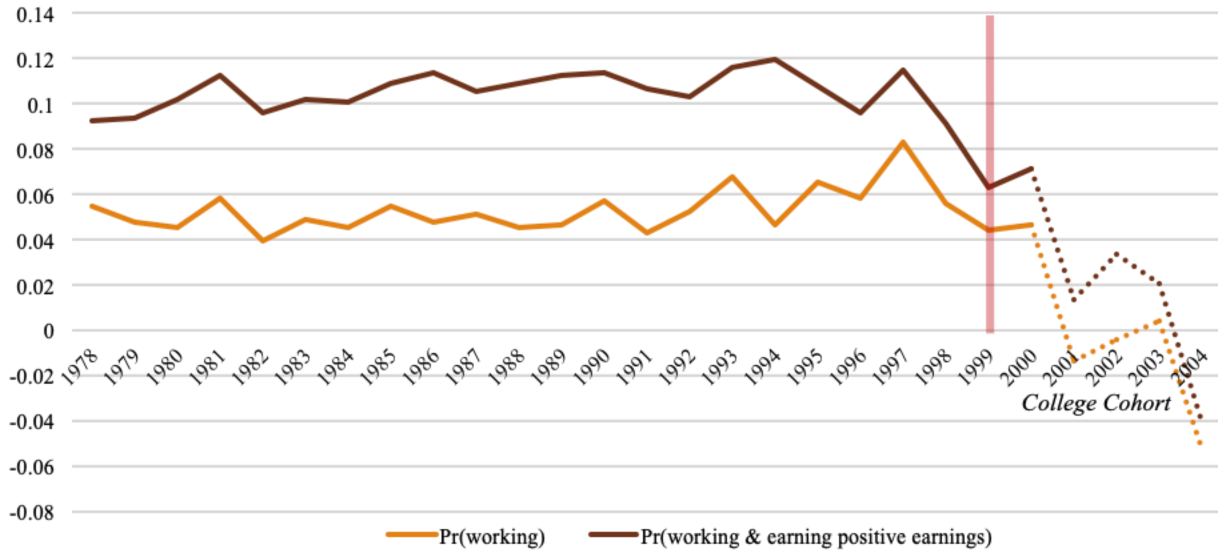
Notes: Annual incomes are discounted to the 2002 Chinese currency (¥) value using annual nationwide CPIs in urban areas obtained from the China Statistical Yearbooks, 2002-2009.

Figure A4: Aggregate Relative Supply Index for BA- vs. HS-educated Labor by Year.



Data source: Author's calculation using data from National Statistics Bureau of China (retrived from: http://www.stats.gov.cn/english/) and various years of *Educational Statistics Yearbook of China.*

Figure A5: BA-HS Gaps in the Probabilities of Having a Job and Earning Positive Earnings by Cohort.



Data source: Urban Household Survey, 2002-2009.

Table A1: Undergraduate Enrollment, Graduates, and Number of Regular Degree-granting HEIs in China (in millions), 1995-2018.

Year	Admission quota for 1st-time-1st-year undergraduate students in regular HEIs (million)	Actual enrollment of 1st-time-1st-year undergraduate students in regular HEIs (million)	Change in new enrollment from the previous year (%)	Total undergraduate enrollment in regular HEIs (million)	College graduates from regular HEIs (million)	Number of regular HEIs
1995	0.93	0.93	3.3	2.91	0.81	1,054
1996	0.97	0.97	4.3	3.02	0.84	1,032
1997	1.00	1.00	3.1	3.17	0.83	1,020
1998	1.08	1.08	8.0	3.41	0.83	1,022
1999	1.55	1.60	48.1	4.13	0.85	1,071
2000	2.21	2.21	38.1	5.56	0.95	1,041
2001	2.68	2.68	21.3	7.19	1.04	1,225
2002	3.20	3.21	19.8	9.03	1.34	1,396
2003	3.82	3.82	19.0	11.09	1.88	1,552
2004	4.47	4.47	17.0	13.33	2.39	1,731
2005	5.04	5.05	13.0	15.62	3.07	1,792
2006	5.62	5.46	8.1	17.39	3.78	1,867
2007	5.66	5.66	3.7	18.85	4.48	1,908
2008	6.08	6.08	7.4	20.21	5.12	2,263
2009	6.39	6.40	5.3	21.45	5.31	2,305
2010	6.62	6.62	3.4	22.32	5.75	2,358
2011	---	6.82	3.0	23.09	6.08	2,409
2012	---	6.89	1.0	23.91	6.25	2,442
2013	---	7.00	1.6	24.68	6.39	2,491
2014	---	7.21	3.0	25.48	6.59	2,529
2015	---	7.38	2.4	26.25	6.81	2,560
2016	---	7.49	1.5	26.96	7.04	2,596
2017	---	7.62	1.7	27.54	7.36	2,631
2018	---	7.91	3.8	28.31	7.50	2,663

Data source: National Statistics Bureau of China (retrived from: http://www.stats.gov.cn/english/).

Table A2: Province-year Level College Admission Ratio, Expansion Intensity, and Classification of High- or Low-intensity Groups, 1995-2008.

College Cohort /Province *	1995	1996	1997	1998	1999			2000			2001		
	admission ratio	admission ratio	admission ratio	admission ratio	admission ratio	intensity (change from baseline)	higher/ lower than nationwide intensity	admission ratio	intensity (change from baseline)	higher/ lower than nationwide intensity	admission ratio	intensity (change from baseline)	higher/ lower than nationwide intensity
Nationwide	**0.067**	**0.065**	**0.067**	**0.056**	**0.082**	**0.026**	---	**0.113**	**0.057**	---	**0.133**	**0.077**	---
Beijing	0.407	0.401	0.412	0.350	0.398	0.078	H	0.424	0.104	H	0.502	0.183	H
Tianjin	0.246	0.238	0.232	0.196	0.260	0.064	H	0.327	0.131	H	0.356	0.160	H
Hebei	0.048	0.045	0.048	0.051	0.072	0.021	L	0.103	0.052	L	0.125	0.074	L
Shanxi	0.067	0.064	0.065	0.050	0.076	0.027	H	0.093	0.043	L	0.111	0.062	L
Inner Mongolia	0.054	0.053	0.055	0.037	0.048	0.011	L	0.081	0.044	L	0.098	0.061	L
Liaoning	0.133	0.130	0.132	0.088	0.139	0.051	H	0.178	0.090	H	0.190	0.102	H
Jilin	0.116	0.114	0.113	0.093	0.119	0.039	H	0.170	0.090	H	0.177	0.096	H
Heilongjiang	0.097	0.094	0.093	0.064	0.105	0.041	H	0.131	0.067	H	0.154	0.090	H
Shanghai	0.392	0.365	0.376	0.296	0.370	0.081	H	0.355	0.066	H	0.437	0.148	H
Jiangsu	0.086	0.090	0.094	0.081	0.126	0.045	H	0.168	0.088	H	0.216	0.135	H
Zhejiang	0.074	0.076	0.076	0.058	0.090	0.032	H	0.119	0.061	H	0.174	0.116	H
Anhui	0.045	0.044	0.046	0.034	0.050	0.016	L	0.080	0.046	L	0.096	0.062	L
Fujian	0.064	0.064	0.068	0.057	0.077	0.020	L	0.096	0.039	L	0.096	0.039	L
Jiangxi	0.059	0.058	0.058	0.037	0.063	0.025	L	0.087	0.050	L	0.111	0.074	L
Shandong	0.050	0.048	0.048	0.044	0.058	0.014	L	0.090	0.046	L	0.121	0.077	H
Henan	0.035	0.034	0.035	0.032	0.049	0.018	L	0.075	0.044	L	0.085	0.053	L
Hubei	0.075	0.074	0.075	0.067	0.101	0.034	H	0.154	0.087	H	0.173	0.106	H
Hunan	0.053	0.052	0.055	0.051	0.081	0.030	H	0.105	0.054	L	0.111	0.061	L
Guangdong	0.051	0.054	0.055	0.053	0.069	0.017	L	0.095	0.042	L	0.101	0.048	L
Guangxi	0.039	0.038	0.041	0.034	0.042	0.008	L	0.064	0.030	L	0.074	0.040	L
Hainan	0.043	0.043	0.043	0.043	0.045	0.002	L	0.066	0.023	L	0.080	0.037	L
Chongqing **	---	---	---	0.047	0.069	0.023	L	0.092	0.045	L	0.110	0.063	L
Sichuan	0.058	0.057	0.040	0.033	0.050	0.016	L	0.070	0.036	L	0.081	0.047	L
Guizhou	0.024	0.024	0.026	0.028	0.050	0.022	L	0.070	0.041	L	0.089	0.061	L
Yunnan	0.038	0.038	0.040	0.035	0.049	0.014	L	0.063	0.028	L	0.078	0.043	L
Xizang	0.027	0.013	0.016	0.118	0.137	0.019	L	0.189	0.070	H	0.172	0.054	L
Shaanxi	0.084	0.084	0.082	0.076	0.146	0.050	H	0.201	0.105	H	0.220	0.124	H
Gansu	0.055	0.054	0.056	0.053	0.073	0.019	L	0.106	0.052	L	0.138	0.084	H
Qinghai	0.051	0.051	0.054	0.043	0.051	0.007	L	0.103	0.060	H	0.121	0.078	H
Ningxia	0.048	0.048	0.050	0.037	0.044	0.007	L	0.076	0.038	L	0.095	0.057	L
Xinjiang	0.059	0.056	0.056	0.048	0.077	0.029	H	0.130	0.082	H	0.157	0.108	H

(continued on next page)

Table A2: Province-year Level College Admission Ratio, Expansion Intensity, and Classification of High- or Low-intensity Groups, 1995-2008. (Cont'd)

College Cohort /Province *	2002 admission ratio	2002 intensity (change from baseline)	2002 higher/ lower than nationwide intensity	2003 admission ratio	2003 intensity (change from baseline)	2003 higher/ lower than nationwide intensity	2004 admission ratio	2004 intensity (change from baseline)	2004 higher/ lower than nationwide intensity	2005 admission ratio	2005 intensity (change from baseline)	2005 higher/ lower than nationwide intensity
Nationwide	**0.162**	**0.106**	---	**0.195**	**0.139**	---	**0.211**	**0.155**	---	**0.218**	**0.162**	---
Beijing	0.631	0.311	H	0.778	0.458	H	0.740	0.420	H	0.724	0.404	H
Tianjin	0.474	0.278	H	0.641	0.445	H	0.544	0.348	H	0.658	0.462	H
Hebei	0.141	0.090	L	0.164	0.113	L	0.181	0.129	L	0.174	0.123	L
Shanxi	0.137	0.087	L	0.161	0.112	L	0.187	0.137	L	0.208	0.159	L
Inner Mongolia	0.109	0.071	L	0.146	0.109	L	0.168	0.131	L	0.167	0.130	L
Liaoning	0.262	0.175	H	0.328	0.240	H	0.332	0.244	H	0.313	0.225	H
Jilin	0.227	0.147	H	0.280	0.200	H	0.274	0.194	H	0.292	0.211	H
Heilongjiang	0.192	0.127	H	0.207	0.143	H	0.201	0.137	H	0.208	0.144	L
Shanghai	0.541	0.252	H	0.679	0.390	H	0.677	0.388	H	0.638	0.349	H
Jiangsu	0.244	0.164	H	0.292	0.212	H	0.330	0.249	H	0.340	0.259	H
Zhejiang	0.242	0.184	H	0.296	0.237	H	0.311	0.253	H	0.296	0.238	H
Anhui	0.119	0.084	L	0.133	0.099	L	0.158	0.124	L	0.168	0.134	L
Fujian	0.108	0.051	L	0.160	0.103	L	0.177	0.120	L	0.213	0.156	L
Jiangxi	0.146	0.108	H	0.188	0.151	H	0.240	0.203	H	0.254	0.217	H
Shandong	0.146	0.102	L	0.179	0.135	L	0.192	0.148	L	0.213	0.169	H
Henan	0.103	0.071	L	0.115	0.084	L	0.145	0.114	L	0.138	0.107	L
Hubei	0.227	0.160	H	0.259	0.192	H	0.282	0.214	H	0.280	0.213	H
Hunan	0.151	0.100	L	0.184	0.133	L	0.185	0.134	L	0.207	0.157	L
Guangdong	0.119	0.067	L	0.150	0.097	L	0.170	0.117	L	0.189	0.136	L
Guangxi	0.082	0.048	L	0.092	0.059	L	0.114	0.081	L	0.118	0.084	L
Hainan	0.100	0.057	L	0.115	0.072	L	0.132	0.089	L	0.165	0.122	L
Chongqing **	0.147	0.100	L	0.258	0.211	H	0.234	0.187	H	0.212	0.165	H
Sichuan	0.119	0.085	L	0.197	0.164	H	0.193	0.159	H	0.192	0.159	L
Guizhou	0.077	0.049	L	0.097	0.069	L	0.101	0.073	L	0.107	0.079	L
Yunnan	0.089	0.054	L	0.104	0.069	L	0.124	0.089	L	0.132	0.097	L
Xizang	0.219	0.100	L	0.226	0.108	L	0.250	0.132	L	0.254	0.135	L
Shaanxi	0.263	0.167	H	0.273	0.177	H	0.286	0.190	H	0.280	0.184	H
Gansu	0.153	0.100	L	0.164	0.111	L	0.162	0.109	L	0.164	0.110	L
Qinghai	0.124	0.081	L	0.154	0.111	L	0.153	0.110	L	0.174	0.131	L
Ningxia	0.103	0.065	L	0.117	0.079	L	0.139	0.102	L	0.142	0.104	L
Xinjiang	0.153	0.105	L	0.150	0.102	L	0.159	0.111	L	0.160	0.112	L

(continued on next page)

Table A2: Province-year Level College Admission Ratio, Expansion Intensity, and Classification of High- or Low-intensity Groups, 1995-2008. (Cont'd)

College Cohort /Province *	2006			2007			2008		
	admission ratio	intensity (change from baseline)	higher/ lower than nationwide intensity	admission ratio	intensity (change from baseline)	higher/ lower than nationwide intensity	admission ratio	intensity (change from baseline)	higher/ lower than nationwide intensity
Nationwide	**0.232**	**0.176**	---	**0.236**	**0.180**	---	**0.258**	**0.202**	---
Beijing	0.681	0.361	H	0.754	0.435	H	0.798	0.478	H
Tianjin	0.691	0.495	H	0.714	0.518	H	0.828	0.632	H
Hebei	0.201	0.150	L	0.193	0.141	L	0.205	0.153	L
Shanxi	0.234	0.184	H	0.212	0.162	L	0.223	0.173	L
Inner Mongolia	0.192	0.155	L	0.225	0.187	H	0.239	0.202	L
Liaoning	0.306	0.218	H	0.330	0.242	H	0.385	0.297	H
Jilin	0.276	0.196	H	0.303	0.223	H	0.326	0.246	H
Heilongjiang	0.264	0.200	H	0.294	0.230	H	0.348	0.284	H
Shanghai	0.696	0.407	H	0.772	0.483	H	0.847	0.558	H
Jiangsu	0.363	0.282	H	0.326	0.245	H	0.309	0.229	H
Zhejiang	0.335	0.277	H	0.351	0.293	H	0.386	0.328	H
Anhui	0.186	0.151	L	0.196	0.161	L	0.201	0.167	L
Fujian	0.224	0.167	L	0.242	0.185	H	0.260	0.223	H
Jiangxi	0.278	0.241	H	0.273	0.236	H	0.290	0.253	H
Shandong	0.232	0.188	H	0.271	0.197	H	0.329	0.285	H
Henan	0.153	0.121	L	0.156	0.124	L	0.200	0.168	L
Hubei	0.293	0.226	H	0.266	0.199	H	0.287	0.219	H
Hunan	0.207	0.156	L	0.210	0.159	L	0.211	0.160	L
Guangdong	0.204	0.152	L	0.209	0.156	L	0.222	0.169	L
Guangxi	0.132	0.099	L	0.157	0.123	L	0.171	0.137	L
Hainan	0.235	0.192	L	0.258	0.215	L	0.249	0.206	L
Chongqing **	0.243	0.197	H	0.283	0.236	H	0.297	0.250	H
Sichuan	0.182	0.149	L	0.197	0.164	L	0.212	0.179	L
Guizhou	0.102	0.074	L	0.099	0.071	L	0.108	0.079	L
Yunnan	0.136	0.101	L	0.126	0.091	L	0.146	0.111	L
Xizang	0.232	0.113	L	0.196	0.078	L	0.199	0.081	L
Shaanxi	0.274	0.178	H	0.281	0.185	H	0.307	0.211	H
Gansu	0.191	0.138	L	0.187	0.134	L	0.210	0.157	L
Qinghai	0.161	0.118	L	0.144	0.101	L	0.174	0.131	L
Ningxia	0.177	0.140	L	0.176	0.138	L	0.203	0.166	L
Xinjiang	0.163	0.115	L	0.160	0.111	L	0.163	0.115	L

Data source: Author's calculation using administrative data from MoE's Department of Development and Planning.

Notes: The statistics cover all the provinces in mainland China except for Hainan. The Chongqing municipality was established in 1997, so data were missing for the 1995-1997 years for Chongqing.

Table A3: Means of Education Outcomes and Earnings by Cohorts and Program Intensity.

A. Workers with positive earnings only

Outcomes	Before-expansion & Low-intensity	Before-expansion & High-intensity	Post-expansion & Low-intensity	Post-expansion & High-intensity	*Difference in Differences*
Years of schooling	11.2	11.4	12.9	13.7	*0.4*
completed	(2.8)	(2.8)	(2.9)	(2.8)	*(0.2)*
Probability of attending	0.265	0.286	0.508	0.605	*0.075*
college	(0.442)	(0.452)	(0.500)	(0.489)	*(0.025)*
Probability of obtaining	0.233	0.257	0.487	0.572	*0.060*
any postsecondary	(0.423)	(0.437)	(0.500)	(0.495)	*(0.024)*
Log(hourly total income)	2.605	2.751	2.553	2.734	*0.034*
	(0.821)	(0.816)	(0.759)	(0.780)	*(0.044)*

B. Full sample with imputed earnings for non-workers

Outcomes	Before-expansion & Low-intensity	Before-expansion & High-intensity	Post-expansion & Low-intensity	Post-expansion & High-intensity	*Difference in Differences*
Years of schooling	10.9	11.2	12.8	13.5	*0.4*
completed	(2.7)	(2.8)	(2.9)	(2.8)	*(0.1)*
Probability of attending	0.232	0.257	0.484	0.582	*0.073*
college	(0.422)	(0.437)	(0.500)	(0.493)	*(0.021)*
Probability of obtaining	0.202	0.230	0.459	0.548	*0.061*
any postsecondary	(0.402)	(0.421)	(0.499)	(0.498)	*(0.020)*
Probability of working	0.828	0.857	0.849	0.892	*0.014*
and earning positive	(0.377)	(0.350)	(0.358)	(0.311)	*(0.017)*
Log(hourly total income)	2.222	2.418	2.236	2.506	*0.074*
	(1.224)	(1.217)	(1.143)	(1.100)	*(0.057)*

Data source: Chinese Household Income Project 2013.

Table A4: Effects of the Expansion on Highest Educational Levels Completed.

	[1] Full Sample *(N=8,377)*		[2] Workers Sample *(N=7,106)*		[3] Employees Sample *(N=5,702)*	
Outcomes	***Coefficient***		***Coefficient***		***Coefficient***	
Completed lower secondary school only (i.e., not attended high school)	-0.229		-0.226		-0.132	
	(0.156)		(0.164)		(0.167)	
Some high school, but no HS diploma	-0.004		-0.004		-0.005	
	(0.021)		(0.022)		(0.024)	
Vocational high school diploma	-0.011		0.046		0.064	
	(0.120)		(0.130)		(0.140)	
Academic high school diploma	-0.366	*	-0.296	~	-0.279	~
	(0.144)		(0.152)		(0.158)	
Any HS diploma	-0.375	*	-0.248		-0.215	
	(0.167)		(0.178)		(0.187)	
Some college, but no post-secondary degree	0.095		0.112	~	0.115	~
	(0.059)		(0.064)		(0.067)	
Associate's degree	-0.161		-0.209	~	-0.262	~
	(0.111)		(0.123)		(0.134)	
Bachelor's degree	0.622	***	0.535	***	0.463	**
	(0.120)		(0.135)		(0.150)	
Graduate degrees	0.151	**	0.149	**	0.151	*
	(0.051)		(0.058)		(0.064)	
Bachelor's degree or above	0.767	***	0.677	***	0.614	***
	(0.126)		(0.141)		(0.156)	

*~ p<.1; * p<.05; ** p<.01; *** p<.001*

Data source: Chinese Household Income Project 2013.

Notes: Robust standard errors clustered at the province level are in parentheses.

***, **, *, and ~ indicate significance at the 0.1, 1, 5, and 10 percent levels.

Table A5: Number of Regular HEIs, Undergraduate Enrollment and Student-faculty Ratio in Regular HEIs in China, 1994-2018.

Year	# of regular HEIs	New enrollment of undergraduate students in regular HEIs (million)	Change from the previous year (%)	Total enrollment in regular HEIs (million)	Average enrollment per regular HEI	Average student-faculty ratio in regular HEIs
1995	1,054	0.93	3.3	2.91	2,757	8.9
1996	1,032	0.97	4.3	3.02	2,927	9.6
1997	1,020	1.00	3.1	3.17	3,122	9.8
1998	1,022	1.08	8.0	3.41	3,315	11.6
1999	1,071	1.60	48.1	4.13	3,815	13.4
2000	1,041	2.21	38.1	5.56	5,289	16.3
2001	1,225	2.68	21.3	7.19	5,870	18.2
2002	1,396	3.21	19.8	9.03	6,471	19.0
2003	1,552	3.82	19.0	11.09	7,143	17.0
2004	1,731	4.47	17.0	13.33	7,704	16.2
2005	1,792	5.05	13.0	15.62	8,717	16.9
2006	1,867	5.46	8.1	17.39	9,314	18.0
2007	1,908	5.66	3.7	18.85	9,879	17.3
2008	2,263	6.08	7.4	20.21	8,931	17.2
2009	2,305	6.40	5.3	21.45	9,306	17.3
2010	2,358	6.62	3.4	22.32	9,466	17.3
2011	2,409	6.82	3.0	23.09	9,585	17.4
2012	2,442	6.89	1.0	23.91	9,791	17.5
2013	2,491	7.00	1.6	24.68	9,908	17.5
2014	2,529	7.21	3.0	25.48	10,075	17.7
2015	2,560	7.38	2.4	26.25	10,254	17.7
2016	2,596	7.49	1.5	26.96	10,385	17.1
2017	2,631	7.62	1.7	27.54	10,468	17.5
2018	2,663	7.91	3.8	28.31	10,631	17.6

Data source: Author's calculation using data from National Statistics Bureau of China (retrived from: http://www.stats.gov.cn/english/).

Table A6: Mean Log Annual Job Income by Age-group and Year for HS- and BA-educated Workers, and the Corresponding BA-HS Earnings Gap.

	Survey year							
Age group	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**
1. HS-educated labor								
Workers Sample	8.939	8.948	9.014	9.080	9.193	9.310	9.441	9.608
(Age 16-49)	(0.766)	(0.807)	(0.789)	(0.801)	(0.777)	(0.742)	(0.830)	(0.711)
Age 24-27	8.742	8.735	8.804	8.850	9.001	9.119	9.264	9.359
	(0.721)	(0.784)	(0.780)	(0.774)	(0.742)	(0.670)	(0.790)	(0.733)
Age 28-31	8.778	8.801	8.922	8.970	9.086	9.210	9.387	9.524
	(0.765)	(0.814)	(0.781)	(0.761)	(0.763)	(0.734)	(0.792)	(0.701)
Age 32-35	8.958	8.962	9.032	9.153	9.205	9.285	9.471	9.613
	(0.781)	(0.830)	(0.793)	(0.776)	(0.775)	(0.744)	(0.781)	(0.690)
Age 36-39	9.007	9.027	9.056	9.134	9.259	9.372	9.482	9.657
	(0.751)	(0.760)	(0.789)	(0.798)	(0.772)	(0.745)	(0.864)	(0.719)
Age 40-43	9.042	9.059	9.097	9.149	9.261	9.376	9.481	9.678
	(0.760)	(0.800)	(0.776)	(0.833)	(0.788)	(0.747)	(0.860)	(0.696)
2. BA-educated labor								
Workers Sample	9.550	9.568	9.680	9.710	9.787	9.890	10.031	10.187
(Age 16-49)	(0.590)	(0.753)	(0.712)	(0.728)	(0.740)	(0.695)	(0.758)	(0.692)
Age 24-27	9.249	9.170	9.258	9.299	9.339	9.502	9.650	9.749
	(0.750)	(0.952)	(0.799)	(0.832)	(0.866)	(0.763)	(0.806)	(0.772)
Age 28-31	9.454	9.477	9.571	9.623	9.727	9.845	9.987	10.131
	(0.544)	(0.766)	(0.741)	(0.752)	(0.658)	(0.659)	(0.735)	(0.662)
Age 32-35	9.558	9.599	9.719	9.722	9.827	9.922	10.125	10.259
	(0.593)	(0.687)	(0.736)	(0.674)	(0.662)	(0.595)	(0.731)	(0.699)
Age 36-39	9.605	9.673	9.762	9.847	9.902	10.018	10.155	10.304
	(0.546)	(0.616)	(0.646)	(0.653)	(0.713)	(0.642)	(0.736)	(0.615)
Age 40-43	9.665	9.747	9.819	9.883	9.957	10.089	10.221	10.365
	(0.524)	(0.659)	(0.595)	(0.652)	(0.690)	(0.676)	(0.644)	(0.573)
3. BA-HS wage gap								
Age 16-49	0.611	0.620	0.666	0.629	0.594	0.580	0.590	0.579
Age 24-27	0.507	0.435	0.454	0.449	0.338	0.383	0.386	0.390
Age 28-31	0.676	0.676	0.649	0.652	0.641	0.635	0.600	0.607
Age 32-35	0.600	0.637	0.687	0.569	0.622	0.637	0.655	0.646
Age 36-39	0.598	0.646	0.706	0.713	0.643	0.646	0.673	0.647
Age 40-43	0.623	0.688	0.722	0.734	0.696	0.713	0.740	0.687

Data source: Urban Household Survey, 2002-2009.

Notes: Annual incomes are discounted to the 2002 Chinese currency (¥) value using annual nationwide CPIs in urban areas obtained from the China Statistical Yearbooks, 2002-2009.

Table A7: Aggregate Relative Supply Index for BA- vs. HS-educated Labor by Year.

	Survey year							
Measures for aggregate relative supply	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**
Naïve aggregate relative supply, age 24-39	0.086	0.094	0.097	0.102	0.110	0.120	0.132	0.146
Naïve aggregate relative supply, age 16-49	0.072	0.079	0.086	0.095	0.106	0.117	0.128	0.138
Aggregate supply with imperfect substitution, age 24-39, workers only	0.098	0.106	0.109	0.114	0.121	0.132	0.143	0.157
Aggregate supply with imperfect substitution, age 16-49, workers only	0.081	0.088	0.095	0.106	0.118	0.129	0.139	0.149
Aggregate supply with imperfect substitution, age 24-39, full sample	0.131	0.139	0.141	0.147	0.156	0.168	0.180	0.197
Aggregate supply with imperfect substitution,age 16-49, full sample	0.105	0.111	0.118	0.132	0.147	0.161	0.169	0.179

Data source: Author's calculation using data from National Statistics Bureau of China (retrived from: http://www.stats.gov.cn/english/) and various years of *Educational Statistics Yearbook of China.*

Appendix B: Additional Notes on Robustness Checks for Chapter 2

I first regress province-year level expansion intensity for the "post" cohorts on a serious of independent variables that might determine the intensity and at the same time would have direct effects on educational and labor market outcomes. The independent variables I test for include province-year level time-varying variables that capture post-expansion trends in: per capita GDP growth, dummies for two influential concurrent reforms (China's accession into the World Trade Organization and the reforms of state-owned enterprises), and the province-level number of HEIs at the baseline level (i.e., in 1998 prior to the expansion). I explain each of these variables in more details below.

Per-capita GDP. I test whether the expansion intensity is correlated with both the current and expected future economic conditions at the province level. More specifically, I construct measures of the annual growth rate of province level per-capita GDP for the current year, previous year, and in four years to come using data from the Bureaus of Statistics, adjusted by CPIs across years. For example, for an individual from probable college cohort of 2008 and from province p, the three GDP measures correspond to the $\frac{per-capita\ GDP_{p,2008}}{per-capita\ GDP_{p,2007}}$ - 1, $\frac{per-capita\ GDP_{p,2007}}{per-capita\ GDP_{p,2006}}$ - 1, and $\frac{per-capita\ GDP_{p,2012}}{per-capita\ GDP_{p,2011}} - 1$. For the two variables measuring the current economic condition, the annual growth rate of per-capita GDP for the previous year is a better predictor of intensity.[39] As for the last GDP measure, since the validity of my instrument lies on the assumption that variation of the intensity does not depend on expected economic growth (otherwise, the estimates of the expansion's earnings effects would be biased upwardly). To test

[39] Also, college admission quota is always designed and announced in early months of each year (i.e., January or February), so it is also make sense that previous year's GDP growth is more correlated with the "intensity" measure than the current year's.

that, I created a "lagged GDP growth rate" variable ($GDP\_lag_{p,c}$) equals to the annual growth rate of GDP for province p in the year that is 4 years after the individual i's college cohort c. This could be viewed as measures for the economic condition when the individual graduate from college and about to enter the labor market. In the regression of determines for the intensity measures for the post cohorts, coefficients on this variable is not statistically significant at the 10 percent level. Combined, these results suggest that the college expansion intensity is at most responsive to the current economic conditions, not a proactive action to expected economic development in the future (or it is hard to correctly expect economic development in the future during a period of rapid and unpredictable development in China). Therefore, I later only include the "annual growth rate of per-capita GDP for previous year" variable as covariates in equation (2.5).

Accession into the WTO in late 2001. China's accession to the WTO in November of 2001 led to accelerated economic openness and institutional improvement (Branstetter & Lardy, 2008). In particular, it disproportionately benefits the import/export industry due to fewer trade barriers and lower tariffs after the accession. If provinces that have more import/export-intensive industry are also more likely to have higher intensity of college expansion, my estimates of the expansion's effects will be biased upwardly. To mitigate the confounding effects due to the potential correlation between the provincial variation in how much they benefit from the accession into WTO and expansion intensity, I control for interactions between a dummy indicating whether the province has a share of GDP from import/export out of the total provincial GDP higher than the national mean for each of the post-2001 years and the corresponding year dummies (in fact, whether the province is import/export-intensive is constant during 2002-2008, i.e., the same set of nine provinces are above the national mean of the share of

GDP from the import/export industry for all the seven years, so I only include the interaction between one dummy for import/export-intensive provinces and one dummy for post-2001 years in the regression).

SOE reforms. As China transits from a centrally-planned economy to a more market-oriented ones starting from December of 1978, there have been gradual reforms of the state-owned enterprises. The most radical period of reforms happened during 1993 to 2001 that altered the ownership of many SOEs and caused a substantial number of laid-off workers from such institutions. As pointed out in section 5 of chapter 1, one of the motivations for college expansion was to induce more high school graduates into colleges to ease the unemployment issues due to an increasing number of laid-off workers because of the SOEs. Given this background, one may suspect that provinces with more severe unemployment problems would be more willing to expand their college enrollment. For this reason, I compute province-year level unemployment rate using data from the Bureau of Statistics for the years during 1998 to 2001. I experimented a couple of measures of the unemployment issue due to SOEs, including using the exact number of unemployment rates at the province-year level, dummies for whether the unemployment rate is above the national average for each of the years during 1998 to 2001, and one dummy indicating whether the mean level of unemployment rates for the years between 1998 to 2001 for province p is higher than the corresponding national average during the same time period. The last measure is the best one, and I use it in the intensity determinants regression and later add it as a covariate in equation (2.5).

Number of HEIs prior to the expansion. The expansion was achieved mainly by expanding enrollment in existing HEIs during initial years of the expansion, and new HEIs have been built in more recent years after the expansion. Therefore, I use the number of HEIs in 1998 at the province level to measure provinces' capacity to expand college enrollment, and this measure is positively correlated with the expansion "intensity" measure during my sample period.