ESSAYS ON DEMOGRAPHIC CHANGE AND EDUCATION

BY HAN LIU

A dissertation submitted to the

Graduate School—New Brunswick

Rutgers, The State University of New Jersey

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

Graduate Program in Economics

Written under the direction of

Mark Killingsworth

and approved by

New Brunswick, New Jersey

October, 2017

ABSTRACT OF THE DISSERTATION

Essays on demographic change and education

by Han Liu

Dissertation Director: Mark Killingsworth

This dissertation discusses a wide variety of topics on demographic change and education. In chapter 2, I examin the fertility effect of Hurricane Katrina. After Hurricane Katrina in 2005, official birth records showed a significant increase in the fertility rate in hurricane destroyed areas. After carefully examining data quality and possible problems with the difference-in-difference method, I conclude that Hurricane Katrina had no significant impact on fertility rate in the affected states. A county-level analysis also supports this conclusion.

In Chapter 3, I study the determinants of the supply of secondary education in China. I consider the unique characteristics of the Chinese educational system: strict hierarchies and consequent competition among schools. Schools on higher level, i.e. high schools, and schools of higher quality, i.e. key-point schools, are at the top tier of the hierarchical system, able to enjoy better resources and serve more capable students. I suggest that when schools and education authorities make joint decisions on enrollment, they face a tradeoff between quantity and quality, and a school's position on the hierarchy determines its specific weighting of quantity versus quality. Using census and aggregate enrollment data, I find strong evidence that cohort-crowding effect for different schools is consistent with their optimal decisions on enrollment. In general a 10% increase in cohort size reduces middle school enrollment rate by less than 1% and high school enrollment rate by 3-4%. Meanwhile, a

10% increase in cohort size reduces teacher/student ratio in primary school, middle school and high school by 3.9%, 3.7% and 2.5% respectively. The variation in different schools' responses to cohort size suggests that schools on top tier of the hierarchy sets higher premium on quality than quantity. Further analysis shows that factors affecting the weighting of quantity at different schooling levels, including the introduction of *Compulsory Education Law* and the number of high schools per county, also have predicted impacts on enrollment. My findings suggest that the strict hierarchy and disparity in resource allocation are the major reasons of inelastic supply in education and I prove that such a system also leads to intense competition among students and loss in efficiency compared to an egalitarian system.

Chapter 4 analyzes the whole college going process from application to admission to matriculation, taking use of the national representative dataset Education Longitudinal Study of 2002 (ELS:2002) which contains students' full records on transition from high school to college. This paper addresses two main problem in estimation: first, for college's admission decision, I use information revealed from application to proxy student's unobserved ability and use conditional logit model to eliminate college fixed effect, solving the problem of endogeneity. Second, for student's matriculation decision, as each student's complete choice set is accessible, I use conditional logit model and nested logit model to obtain unbiased estimation on factors affecting a student's matriculation decision. The results show that students' application to different types of college has different implications on student abilities; gender, race, SAT score and academic performance in high school affect students' probability to be admitted. Also, factors like tuition and distance have significant impact on students' matriculation decisions.

Acknowledgements

I gratefully thank my advisor, Mark Killingsworth, for his long-term guidance and unfailing support. He taught me how to do research and his kindness, sincerity and patience made me a better person. With deep and comprehensive understanding of the field and endless new ideas and smart solutions, he gave me the freedom to explore on my own, and at the same time the guidance to recover when my steps faltered.

I sincerely thank my committee members, Anne Piehl and Ira Gang. Their insightful comments and thought-provoking questions are so valuable for my research. I am grateful to Jennifer Hunt. She pushed me to think further and her numerous comments and suggestions have greatly improved the manuscript.

I also thank my close friends in Rutgers, especially the Pot-Love Group. Their company and encouragement helped me overcome many crisis situations and made my graduate experience so pleasant. Zhutong and Shuyang must be singled out here: they kept me healthy both emotionally and intellectually. I feel extremely fortunate to have them as my life-long friends and role models.

Finally, I must thank my whole family, parents, grandparents, aunts and uncles, for their support and trust during the long period of my pursuit for PhD degree. I am also indebted to my fiance Chuan. Without him, I cannot finish my dissertation so happily and smoothly.

Dedication

This dissertation is dedicated to my dearest parents. Thank you mom and dad for all your love and support throughout my life.

Table of Contents

List of Tables

Chapter 1

Introduction

This dissertation discusses a wide variety of topics on demographic change and education. The first essay is about the fertility effect of big disasters. Although this phenomenon has been explained by previous literature, none of the literature tried to question the problematic birth count data associated with disaster evacuation. By making remedies on data and methodology issues, my findings cast doubt on whether the effect really exists. The second essay is the first paper to study the supply of public education in China using economic analysis. As the behavior of government and educational administrations is like a black box, changes in cohort size is taken as an exogenous shock to test different objective set by different schools. This essay suggests that China maintains a selective secondary education system mainly because administrations has pressure to make an optimal choice between enrollment and quality. The findings have profound implications for policy makers in education system. The third paper focuses on the whole college going process, taking advantage of a rich dataset containing students comprehensive information from application to admission to matriculation. It helps us to obtain unbiased estimation on determinants of admission and matriculation, including solving unobserved student ability issue by adding variables generated from application and admission result.

The second chapter, titled "***Did Hurricane Katrina really increase the fertility rate***?", studies the fertility effect of Hurricane Katrina, one of the most destructive hurricanes in American history. After its landfall as a Category 3 hurricane on August 29 in southeast Louisiana, Katrina swept over the Gulf Coast states including Louisiana, Mississippi, Alabama, Florida etc. The indirect aftermaths of Katrina have been widespread. One surprising finding is the apparent fertility effect in Hurricane impacted states, Louisiana and Mississippi. According to U.S. Vital Statistics Office, from 2005 to 2006 the average births

per 1000 women increased respectively by 12.8% in Louisiana and 11.7% in Mississippi, significantly exceeding the U.S average fertility increase at 3.5%.

This record seems to have added new evidence to the story that disaster is associated with higher fertility rate. Economists have long been interested in fertility as reproduction turned to be a decision-making process (Becker, Duesenberry & Okun, 1960) in societies with contraceptive knowledge. There are several empirical papers that connect fertility and disasters. One strand is to regard the fertility change as the result of immediate sexual behavior at the very point of exogenous shock (Udry, 1970; Evans, Hu and Zhao, 2010). Another strand of research focuses on the long-term disaster effect (Cohan and Cole, 2002, Rodgers, John, and Coleman, 2005).

In Chapter 2, I first used monthly vitality data in state level and DD estimation method, which are most frequently used in fertility estimating process. As Katrina made the landfall at the end of August, 2005, the earliest effect should appear in May or June in 2006, allowing for an average gestation period of 38 weeks. I use births in 2005 as before group and births in 2006 as after group and see whether there is month-specific treatment effects for the treatment group. The result suggests that month-specific treatment effect is significant only after August, indicating that from 2005 to 2006 the fertility increase in hurricane affected states is greater than that in control group states between August to December. This result suggests that there exists positive impact of hurricane. However, I find that the naive estimation using aggregate vital statistics data as well as simple difference in difference method is problematic on both data and methodology, leading to a bias in estimation.

The 2005 data is problematic as the monthly population counts are imprecise after August due to the Hurricane. A simple fix is to use 2004 observation as the before group rather than 2005. I also applied different regression specifications to 2004 and 2006 data. The result shows that the estimated standard errors are different in different regressions, which is consistent with the theoretical prediction by Bertrand et al. (2004). After using two correction methods proposed by Bertrand et al. (2004), I find that Hurricane Katrina shows no significant impact on birth decision in those affected areas.

Since in the dataset county-level data is only available for counties with population

greater than 100,000, I use county-level analysis as a side evidence. The county-level estimation is consistent with the conclusion reached using state-level data. I used two different methods to proxy for degree of destruction in different counties. The change of fertility in FEMA areas showed no difference from those unaffected areas; also, theres no relation between disaster level and fertility change.

The third chapter, titled "***How Responsive is the Supply of Chinese Education to Demographic Change: an Analysis Based on Hierarchy and Competition***", provide a systematic analysis on secondary education system in China. The educational system in China has expanded rapidly in recent decades. The enrollment rate in junior secondary school has risen from 77% in 1992 to about 100% in 2005 and subsequent years. The enrollment rates in senior secondary education and tertiary education have risen from 26.0% and 3.9% respectively in 1992 to 82.5% and 26.5% respectively in 2010. This expansion is likely to have been an important contributor to the rapid growth in Chinese GDP: the increase in human capital, including more universal access to education, create strong momentum for economic development in China and all around the world (eg. Barro and Lee, 1994; Li and Huang, 2009). In addition to boosting growth, the expansion in the supply of educated workers, viewed in isolation, would be expected to reduce the return to education. In fact, the return to education in China has continued to rise, with the premium for college graduates over high school graduates rising steadily from about 20 log points in 1988 to about 70 log points in 2008 (Han, Liu and Zhang, 2014). This is undoubtedly due to a concomitant rise in the demand for skilled workers caused by the transition from a planned to a market economy (Fleisher and Wang, 2005; Wu and Xie, 2003; Zhou, 2000).

The eagerness of students to acquire more education attests to rising demand. Its widely reported that the majority of third year students in middle school and high school are confined to school for more than ten hours every day to get higher scores in entrance exam to higher schools. In populated cities, parents pay as much as ten thousand dollars per year to procure access to elite public primary and secondary schools (Yang, 2011), for which entrance is otherwise based on vicinity of residence or exam results. Even kindergarten graduates are reported to learn middle school materials to succeed in primary school admission exam. In 2003 the family expenditure on education accounted for 12.6% of the total

family budget in Chinese cities (Yu and Suen, 2004).

It is therefore possible that the supply is still in shortage despite the expansion. To gain insight, I seek in this paper to understand the determinants of the supply of education in China and root causes of intense competition. No existing paper performs this economic and quantitative analysis, to my knowledge, although the intense competition has drawn sharp criticism from the general public and the academia. I consider the unique characteristics of the Chinese educational system: strict hierarchies and consequent competition among schools. Higher level schools, i.e. high school, and schools of higher quality, i.e. key-point schools, are at the top tier of the hierarchical system, able to enjoy better resources and serve more capable students. I begin with a theoretical model, in which I suggest that when schools and education authorities make joint decisions on enrollment, they face a tradeoff between quantity, since both individuals and government's long-term objective of economic development demands high enrollment, and high quality, since school officials need to provide high quality relative to other schools to vie for more government appropriation and attract more talented students. My model predicts that the optimal levels of enrollment and quality for revenue-constrained schools are determined by three major arguments: demand, resources available, and the school's specific weighting of enrollment (w) versus quality (1-w). Generally schools on top tier of the hierarchy set a higher weight on quality. I then provide empirical analysis examining the responsiveness of the supply of education, especially to changes in cohort size, and also test the model by estimating the impact on enrollment and quality of relative weights schools put on enrollment versus school quality.

The benchmark empirical analysis is Chinas secondary schools enrollment elasticity with respect to cohort size, underscoring the mechanism that constrained supply creates crowding-out effect in educational institutions. There are only two papers studying the cohort size effect in mainland China: Coxhead, Shen, and Yao (2014) found that baby boomers born after 1961 benefited from large-scale school construction in 1970s and had higher primary school enrollment rate. Using individual level data, Ma (2014) found that one born in smaller cohorts compared to future cohorts enjoy more years of education. Compared to their reduced-form analysis, my analysis is based on schools optimal choice

on enrollment quota and I find that in general a 10% increase in cohort size reduces middle school enrollment rate by less than 1% and high school enrollment rate by 3-4%. Meanwhile, a 10% increase in cohort size reduces teacher/student ratio in primary school, middle school and high school by 3.9%, 3.7% and 2.5% respectively.

To further evaluate my theory, I consider variations in weight on quantity (w) versus quality ($1 - w$) across time and provinces for the same type of school. Though w cannot be estimated directly without assuming utility function form or having individual school's data, I find variables that may affect w or serve as an indicator of w and examine whether the effect of changes in those variables is consistent with the effect of changes in w as is predicted by the model. In middle school level, since the implementation of *Compulsory Education Law* is essentially a process for the authorities to put higher weight on enrollment than quality, I estimate its impact on enrollment and quality and find that, as predicted, it lowered fat cohorts loss in terms of enrollment opportunities but exacerbated their disadvantage in terms of quality. In high school level I use the number of high schools per county as an indictor for w. I find that one additional high school increases enrollment elasticity with respect to cohort size by 0.1 while reducing quality elasticity with respect to cohort size by .04. I also find that officially designated key-point high schools, who have high criterion on quality, have much lower enrollment elasticities than ordinary high schools. Those findings add evidence on literature about schooling quality (Barro and Lee, 2001; Card and Krueger, 1996; Hanushek, 1992; Harmon and Walker, 2000; Jensen, 2010). Though a set of papers discussed about quality and quantity trade-off, most are from the view of individual's decision making process (see e.g. Guo, Yi and Zhang (2016) for a review). This paper offers a specimen where the government, as a social planner, makes tradeoffs between quantity and quality and shows how educational institutions followed the government policy to realize their specific objectives.

Overall, my results suggest that Chinese schools responsiveness to increased demand for education are consistent with the objectives set by different schools at different time. In such a typical hierarchical system where resource distribution to different schools is extremely uneven, higher weight is put on quality ($1 - w$) for more favored schools, i.e. key-point schools and higher level of schools. On the supply side, this is the reason for rationing

access to schooling above middle school; on the demand side, this is the reason for intense competition among students. The selective and competitive nature revealed in this paper challenges the practice to directly applying traditional Mincers method to measure return to education in China, since China has a totally different education system compared with countries where access to school, at least before tertiary stage, is not constrained by supply[1]. Though a large body of literature has calculated return to education in China, only in a recent paper by Li, Liu and Zhang (2012) the impact of selective educational system was mentioned. They found that compared to other countries in China a very small proportion of return to education (2.7% out of 8.4%) is a result of accumulated human capital and the rest is caused by omitted ability, suggesting non-tertiary education in China serves to pick up the most capable ones rather than teaching students skills used in workplaces. Given the characteristics of Chinas education system, I also develop a signaling model about the selecting process of higher level schools and prove that this process leads to a welfare loss for students.

The fourth chapter, titled "***The College Application, Admission, and Matriculatio Decisions: Applications of Conditional Logit Model***", discusses the whole process of students' transition from high school to college that involves repeated interaction between students and colleges with both parties making series of decisions. In this process, there are at least three stages where the role of decision makers alter: first, students search for information from all sources and discuss with friends, parents and counselors about colleges they have interest in or potential to be admitted. They may choose several colleges strategically to which they send official applications. Second, colleges scrutinize applications submitted by students and decide to admit some of them based on their academic performance, college entrance exam, extracurricular activities, recommendation letter, and personal statement etc. Third, each student compares offers from different colleges and make the final matriculation decision. In this paper, I refer to the three stages as application - admission - matriculation decision respectively, and analyze the separate decision-making processes and the interplay between those processes.

[1]The traditional Mincer function is based on the assumption that the present discounted value of going to school is equal to that of not going to school. Even many papers consider selection and sorting biases, they still assume that individuals have greatest sway in determination of their optimal schooling years.

In economics research, there is numerous literature that involves the college going process where students are selected into different colleges according to their characteristics and budget constraints (for example, Black & Smith, 2004; Dale & Krueger, 2002, 2014). Those papers seek unbiased estimation on how different variables affect one's probability to go to a college. However, research considering the college going process itself is relatively scarce. This chapter fills the gap by analyzing the whole college going process comprised of application, admission and matriculation, taking use of the national representative dataset Education Longitudinal Study of 2002 (ELS:2002) which provides a large variety of variables recording students personal experience from high school to early work stage or further schooling. This paper contributes to existing literature by obtaining unbiased estimation on factors affecting college going process, taking advantage of complete information provided by ELS 2002 about the process from application to matriculation. This paper focuses on two separate decisions: admission decision and matriculation decision. We explicitly treat the two decisions as made by colleges and students separately. Using the conditional logistic model, the methodology in this paper allows for specific choice set faced by all decision makers and we have a large set of student-college matched variables to capture factors considered by each party. Those factors include tuition, distance, and ability matching, etc.

Though we focus on colleges' admission decision and students' matriculation decision, the main contribution is made by adding information generated from application process. When considering the college's admission decision, several papers used the conditional logit model as it can eliminate college-specific fixed effect (for example, Hoxby and Avery, 2004; Klaauw, 2002; Long, 2004). However, it doesn't take care of unobserved student ability. This paper proxies unobserved ability using information revealed in application process by assuming that ability unobservable to econometricians is observable to colleges and students. Therefore students' applications to and admissions by colleges at different selectivity level can tell about the student's ability level. The trick here is that a college cannot observe an applicant's application and admission results at other colleges at all; therefore those variables will not affect the college admission decision directly. If the variables have impact on admission result after controlling for other observed characteristics, it should be through the channel that they are indicators of individuals unobserved abilities. Three sets of

variables about a student's applications to colleges at different selectivity level are used: whether the student applied to colleges at one selectivity level or not, the number of colleges the student applied at one selectivity level, and the student's acceptance rate by colleges at the selectivity level. The result shows that for college at a given level, students who filed more applications to a higher level and less applications to a lower level are more likely to be admitted; higher acceptance rate by other colleges also predict higher probability of admission. The findings here also reveals students' behavior pattern in college application. They typically apply for some "reach" school, some "match" school and some "safety" schools according to perception about their own ability and chances to be admitted by different colleges.

This paper also casts light on college "matching". A burgeoning literature discusses about how low-income or other disadvantaged students select themselves into under-match colleges, i.e. less selective colleges that are below their abilities (Avery & Hoxby, 2013; Cortes & Lincove, 2016; Dillion & Smith, 2017; Smith, Pender, &Howell, 2013). However, the definition of "undermatch" is not very clear. The literature usually focuses on the gap between a student's rank in measurable academic performance and the rank of the college she applied or enrolled and assumes such undermatch is an irrational choice. However, this paper suggests perhaps students apply to undermatched colleges based on their rational expectation about their chances to be admitted. For example, Avery and Hoxby (2013) pointed out that most under-matched high ability students are from districts too small to have selective public high schools and therefore the lack of application is due to lack of information. But our result may suggest another story: due to the lack of exposure to high ability peers or teachers, those students may have disadvantage on other unmeasurable factors like recommendation letter, personal statement etc. and thereby choose not to apply to selective colleges with concerns on their chances to be admitted. Therefore, interventions like more counseling or preparation, rather than simply "informing" or "reaching" those students, might be of greater help.

Chapter 2

Did Hurricane Katrina really increase the fertility rate?

2.1 Introduction

Hurricane Katrina in 2005 was one of the most destructive hurricanes in American history. After its landfall as a Category 3 hurricane on August 29 in southeast Louisiana, Katrina swept over the Gulf Coast states including Louisiana, Mississippi, Alabama, Florida etc. The following storm surge, as well as the hurricane itself, caused severe devastation. The total number of fatalities either directly or indirectly related to Katrina was more than 1,833 (Knabb, Rhome, Brown, 2005). As the costliest hurricane in United States, the estimated economic loss was approximately 108 billion (Blake, Landsea & Gibney, 2011). The scope of human suffering inflicted by Hurricane Katrina in the United States has been greater than that of any hurricane to strike this country in several generations.

The indirect aftermaths of Katrina have been widespread. One surprising finding is the apparent fertility effect in Hurricane impacted states, Louisiana and Mississippi. According to U.S. Vital Statistics Office, from 2005 to 2006 the average births per 1000 women increased respectively by 12.8% in Louisiana and 11.7% in Mississippi, significantly exceeding the U.S average fertility increase at 3.5%.

This record seems to have added new evidence to the story that disaster is associated with higher fertility rate. Folk wisdom believes that people are more likely to have baby when they are seeking peace in their hearts after disaster; tension and the electricity blackout caused by disaster also contribute to trap people to stay at home and engage in more sexual activities. The newspapers have confidently reported small baby boom after terrorist attack of September 11, 2001, New York City blackout of 1965, and more recently, Hurricane Irene of 2011.

Economists have long been interested in fertility as reproduction turned to be a decision-making process (Becker, Duesenberry & Okun, 1960) in societies with contraceptive knowledge. Researchers treat fertility as economic behavior and mainly focus on how birth patterns vary with economic transition or policy change over long term. Fertility behavior during catastrophes - which is relatively independent of social transition is also fascinating for researchers interested in the impact of exogenous shock.

There are several empirical papers that connect fertility and disasters. One strand is to regard the fertility change as the result of immediate sexual behavior at the very point of exogenous shock. Udry (1970) examined the effect of the great blackout in New York City on Nov 10, 1965. Using a simple comparison of birth numbers in that period between year 1966 and previous 5 years, the paper concluded that there was no fertility effect. A recent study by Evans, Hu and Zhao (2010) checked the fertility effect of hurricanes in US. They used storm advisory data as the indicator of hurricane magnitude and assumed storm advisories could influence births by altering peoples sexual activity and contraceptive choice during an advisory. They found that low-severity storm advisories have a significant positive effect on fertility while high-severity advisories have a significant negative effect on births after 9 months. Also, as expected there is weak evidence of fertility effect in the long term since storm advisories are not necessarily associated with real hurricane loss.

Another strand of research focuses on the long-term disaster effect. Cohan and Cole (2002) studied how hurricane Hugo in 1989 impacted fertility, marriage and divorce over a 23-year time span. Using county level data in Southern California, time-series analysis (ARIMA model) showed that in years after the hurricane Hugo, birth rates increased in the counties declared disaster areas compared with the 22 other counties in the state. Another literature paper on the long-term effect of disaster is by Rodgers, John, and Coleman (2005). They studied the fertility effect of the Oklahoma City bombing in 1995 using birth data from 1990 to 1999. Difference-in-difference (DD) method was employed, finding that the overall fertility change after the bombing was significant and positive for Oklahoma County.

In this paper, I first used monthly vitality data in state level and DD estimation method, which are most frequently used in fertility estimating process. As Katrina made the landfall at the end of August, 2005, the earliest effect should appear in May or June in 2006, allowing

for an average gestation period of 38 weeks. I use births in 2005 as before group and births in 2006 as after group and see whether there is month-specific treatment effects for the treatment group. The result suggests that month-specific treatment effect is significant only after August, indicating that from 2005 to 2006 the fertility increase in hurricane affected states is greater than that in control group states between August to December. This result suggests that there exists positive impact of hurricane. However, I find that the naive estimation using aggregate vital statistics data as well as simple difference in difference method is problematic on both data and methodology, leading to a bias in estimation.

The 2005 data is problematic as the monthly population counts are imprecise after August due to the Hurricane. A simple fix is to use 2004 observation as the before group rather than 2005. I also applied different regression specifications to 2004 and 2006 data. The result shows that the estimated standard errors are different in different regressions, which is consistent with the theoretical prediction by Bertrand et al. (2004). After using two correction methods proposed by Bertrand et al. (2004), I find that Hurricane Katrina shows no significant impact on birth decision in those affected areas.

Since in the dataset county-level data is only available for counties with population greater than 100,000, I use county-level analysis as a side evidence. The county-level estimation is consistent with the conclusion reached using state-level data. I used two different methods to proxy for degree of destruction in different counties. The change of fertility in FEMA areas showed no difference from those unaffected areas; also, theres no relation between disaster level and fertility change.

The remainder of the paper is presented in the following way: section 2 introduces methodology and data used. Section 3 is state-level analysis, focusing on the comparison between treatment states and control states. Section 4 is county-level analysis, focusing on whether different levels of hurricane destruction have heterogeneous effect on fertility. Section 5 is robustness check on whether the analysis is valid since hurricane and pregnancy may jointly affect migration decisions. Section 6 is the conclusion.

2.2 Method and Data

Followed Rodgers et al. (2005), I used difference-in-difference (DD) method to capture the fertility effect of the exogenous hurricane shock. DD method is a quasi-experimental technique that measures the effect of a treatment for a given period, frequently used in studies of the impact of exogenous shocks on fertility rate. Using states unaffected by Hurricane Katrina as control group and states devastated as treatment group, DD method can identify the pre-post difference generated by Hurricane Katrina.

One important issue is how to identify control and treatment states. Although Hurricane Katrina spread across several states, the level of devastation varied a lot. A proper way to define each states devastation level is to use loss evaluation by Federal Emergency Management Agency (FEMA). After Katrina, FEMA(2005) quickly evaluated the local destruction and designated many counties throughout the Gulf Coast as "disaster counties". According to FEMA Katrina declaration, all counties in Mississippi and Louisiana, 22 counties in western Alabama, and 11 in Florida were eligible for FEMA assistance. Thus I define Mississippi (MS) and Louisiana (LA) as primary treatment states, Alabama (AL) and Florida (FL) as secondary treatment states. As for control groups, we could use either states adjacent to treatment states (Arkansas, Oklahoma, South Carolina, Tennessee, Texas) or all U.S. states except treatment states. I name them adjacent control states or total control states respectively.

Birth Data used in this paper is collected from the National Vital Statistics System of the National Center for Health Statistics (NCHS). NCHS collects microdata from birth certificates filed by U.S. vital statistics offices in each state. The advantage of NCHS data is that it's based on 100-percent sample of birth certificates which contains demographic and health data including date of birth, mother's age, marital status and race, live-birth order, and geographic area, etc. Although individual' level data is inaccessible for years after 2005, this database offers birth counts by detailed personal characters such as birth month, date, age, mother's education, state of residence, etc. So I group women into several different demographic cells by age, race, state and count the number of births in each month. To focus on fertility behavior, I only consider females in reproductive age (15-44). I categorize

women in reproductive age into 6 groups: 15-19 year-old, 20-24 year-old, 25-29 year-old, 30-34 year-old, 35-39 year old, and 40-44 year-old. The dependent variable of interest, fertility rate, is calculated as birth counts divided by female population in a specific cell.

The U.S. Census Bureau offers official population estimates by state, age, race etc on July 1st for each calendar year. One problem, however, is that the birth count data is for each month while population is for each year. To estimate the monthly population, Rogers et al. (2005) assumed a linear interpolation process. Here I assume a constant rate of monthly population change for a demographic group from July 1st of one specific year to July, 1st of the next year. For example, if the population of white women aged 15-19 in Alabama is X on July 1st 2004 and Y on July 1st 2005, then the population estimation in January 2005 is X*(Y/X)$\hat{(}$6/12).

2.3 State-level Analysis

2.3.1 A simple comparison of 2005 and 2006 data

A natural way to use difference in difference method is to compare fertility between 2005 and 2006. As Katrina made the landfall at the end of August, 2005, we may expect the earliest effect appears at May or June in 2006, allowing for an average gestation period of 38 weeks.

To understand effect of Hurricane Katrina on fertility, I estimate the following regression equation:

$$fer_{ijtm} = \beta_0 + \beta_1 M_m + \beta_2 Y_t + \beta_3 T_i + \gamma_m M_m (Y_t * T_i) + \beta_4 X_j + \beta_5 S_j + \epsilon \tag{2.1}$$

where fer_{ijtm} represent the fertility rate, the total number of newborns by females in state i in year t in month m in age cell j divided by female population. Y_t is an indicator of year effect. Y_t equals 0 for year 2005 and 1 for year 2006. T_i is an indicator of treatment which equals to 1 for observations in treatment states and equals to 0 otherwise. As birth decision reflects seasonal variations, I also include indicator M_m to control for different months. γ_m represents the treatment effect for one specific month. X_j are dummies that

control for the six different age group. S_i is state dummies.

An important design in the regression is the use of monthly indicator to capture the variation of treatment effect in different month. As Hurricane Katrina made a landfall at the end of August, we should look at the result after 9 months to allow for gestation period. The natural speculation is that, if Hurricane Katrina did have effect on fertility rate, the treatment effect would appear after May. Meanwhile, if the treatment effect lasted for several months, the change of fertility rate might be households rational decision in response to large exogenous shock as opposed to transitory deviation caused by black-out effect proposed by Evans et al. (2010).

Table 1 shows the result for the regression. Column 1 uses regression on primary treatment states and adjacent control states. The U.S. total fertility rate increased from 2005 to 2006. The impact of Hurricane Katrina turned out to be significant only from August to December. Compared to 2005, in August 2006 the treatment effect is that the number of monthly births per 1000 women increased by approximately 0.74. Compared with U.S average monthly births per 1000 women at 5.7, this result indicates great fertility effect possibly caused by Katrina. Regression Column 2 added all other U.S. states into the control group. The result shows that compared with states unaffected by Katrina, the number of monthly births in Louisiana and Mississippi increased by about 0.83 per 1000 women from August to December. Column 3 and column 4 used all 4 FEMA states as treatment states. The introduction of the two secondary treatment states lowered the scale of treatment effect, suggesting that the fertility effect in two less affected states, Alabama and Florida, is not as significant as that in Louisiana and Mississippi.

2.3.2 Issues on data quality due to Hurricane Katrina

One puzzle is why the treatment effect manifests only after August 2006. If Hurricane Katrina had immediate impact on fertility behavior, the treatment effect should be significant after May in 2006. The possible rationale is that the result shown in table 1 is invalid due to data quality problem. In normal situations, the NCHS vital statistics records are reliable as the data is directly collected from official birth certificates. However, for extreme situations like Hurricane Katrina, the data collected should be used with super carefulness.

Table 2.1: Estimation on fertility effect by month: 2005 2006

	I	II	III	IV
Dependent variable: fertility rate*1000				
Treatment states:				
Primary	X	X	X	X
Secondary			X	X
Control states:				
Adjacent	X		X	
Overall		X		X
Year2006	0.2049***	0.153***	0.2049***	0.153***
	(4.32)	(4.85)	(4.11)	(4.90)
Treatment	-0.3261***	-0.0999	-0.3126***	-0.0864
	(-5.2)	(-0.9)	(-5.91)	(-1.1)
Jan*Y2006	0.0213	0.2588	-0.00588	0.223
	(0.1)	(0.66)	(-0.04)	(0.8)
Feb*Y2006	0.0143	0.073	0.0297	0.0903
	-0.07	-0.19	-0.18	-0.32
Mar*Y2006	-0.038	-0.0935	-0.0306	-0.0924
	(-0.18)	(-0.24)	(-0.18)	(-0.33)
Apr*Y2006	-0.0554	-0.2621	-0.0901	-0.283
	(-0.26)	(-0.66)	(-0.54)	(-1.01)
May*Y2006	-0.022	-0.1809	-0.0746	-0.2418
	(-0.11)	(-0.46)	(-0.45)	(-0.87)
Jun*Y2006	0.0927	0.0242	-0.0733	-0.1418
	(0.44)	(0.06)	(-0.44)	(-0.51)
Jul*Y2006	0.3204	0.297	0.1257	0.0957
	(1.53)	(0.75)	(0.76)	(0.34)
Aug*Y2006	0.7417***	0.8226**	0.4345***	0.5109*
	(3.54)	(2.08)	(2.62)	(1.83)
Sep*Y2006	0.6486***	0.8053**	0.3081*	0.4716*
	(3.09)	(2.04)	(1.85)	(1.69)
Oct*Y2006	0.7177***	0.8419**	0.4593***	0.5771**
	(3.42)	(2.13)	(2.76)	(2.07)
Nov*Y2006	0.7024***	0.8917**	0.4201**	0.6282**
	(3.35)	(2.26)	(2.53)	(2.25)
Dec*Y2006	0.4769**	0.7657*	0.2587	0.5467*
	(2.28)	(1.94)	(1.56)	(1.96)
agegroup	yes	yes	Yes	yes
state	yes	yes	Yes	yes
month	yes	yes	Yes	yes
R-square	0.9714	0.8656	0.9668	0.8681
N	1008	7056	1296	7344

Table 2 shows the 2005 birth count by place of residence by month of birth in Louisiana and Mississippi. There was a sudden decrease of birth count for some areas after August 2005. For example, Orleans Parish in Louisiana had relatively smooth birth count from January to July (Mean = 524, SD = 34). Starting from August, it decreased by more than 20% in one month and reached the lowest point at 109 in November. The sharp decrease is due to the evacuation after Katrina landfall and would be a great source of bias in our estimation: the great treatment effect after August might be stemmed from low fertility rate in 2005 instead of high fertility rate in 2006. Thus, I suggest to use the birth count in 2004, for which year the official birth certificate result is reliable, as before group.

Table 2.2: Birth count: Louisiana and Mississippi, 2005

DOB_MM	**Jan**	**Feb**	**Mar**	**Apr**	**May**	**Jun**	**Jul**	**Aug**	**Sep**	**Oct**	**Nov**	**Dec**
Louisiana	5361	4866	5385	4892	5200	5399	5109	5242	5186	4766	4680	4851
Caddo	308	296	296	275	299	310	335	319	4766	297	281	292
Calcasieu	220	208	209	186	236	248	206	252	226	206	236	217
East Baton Rouge	456	418	464	419	448	475	435	484	544	496	519	537
Jefferson	525	480	541	459	526	487	467	449	342	352	365	362
Lafayette	243	190	259	234	242	242	238	249	296	250	252	220
Orleans	558	474	572	494	524	530	518	397	193	126	109	128
Ouachita	179	161	197	159	202	195	172	214	189	188	197	206
Rapides	165	150	143	156	149	158	140	164	196	161	147	186
St. Tammany	222	224	257	225	192	254	215	241	224	232	188	208
Tangipahoa	133	138	153	120	129	142	133	132	156	137	115	152
Terrebonne	142	110	127	131	121	132	135	118	152	138	148	125
others	2210	2017	2167	2034	2132	2226	2115	2223	2349	2183	2123	2218
Mississippi	3535	3140	3475	3240	3474	3539	3560	3757	3762	3611	3523	3779
DeSoto	159	117	194	171	179	192	177	155	166	184	175	159
Harrison	240	232	242	207	218	253	273	265	208	216	196	199
Hinds	329	288	320	283	330	317	350	351	360	350	349	349
Jackson	154	136	165	115	145	146	146	144	126	136	140	158
Rankin	194	144	160	157	160	178	159	153	154	171	177	162
others	2459	2223	2394	2307	2442	2453	2455	2689	2748	2554	2486	2752

In addition, it would be problematic if we use annual Census Bureau population estimate on July 1st to estimate monthly population in Louisiana and Mississippi because there was a large flow of evacuation after Katrina landfall. The population estimate would be inaccurate and lead to great bias to study the change of fertility behavior. For data from January to July in 2006, I assume the devastated areas were recovering and there was a steady change of population from January to July. The January 2006 population estimation for FEMA areas can be acquired from American Community Survey (ACS) data. ACS launched a special survey for impacted counties in the gulf coast area, which gives population estimation on Jan.1st, 2006. I use this dataset to estimate the monthly population from January to June by assuming a constant rate of population change in Katrina affected areas.

Table 3 shows the result using 2004 and 2006 birth data and revised population estimate. The model specification is the same as the regression exhibited in table 1. The coefficient of month*year2006 shows that the corresponding impact on fertility decreased as predicted. However, the treatment effect is still highly significant in August.

2.3.3 Issues on Methodology

A problem concerning difference-in-difference method is serial correlation, as is proposed by Bertrand, Duflo and Mullainathan (2002). They argue that DD method would underestimate the standard error and thus reject the null hypothesis at a falsely high rate even for a placebo intervention; estimation using single prior- post period would be more reliable.

The estimations mentioned in previous section, as well as Rogers et.al (2005) paper using DD method, rely on data for several periods. Thus the estimation could be highly biased due to serial correlation problem. A Wald test on previous estimation rejects the hypothesis that theres no first-order autocorrelation at 1% level (F = 21.5).

To examine to which extent the estimation is biased, I run regressions on different model specifications.

$$fer_{ijt} = \beta_0 + \beta_1 Y_t + \beta_2 T_i + \gamma(Y_t * T_i) + \beta_3 X_j + \epsilon \tag{2.2}$$

As a comparison to estimation in previous sections, I divided all observations by month and check the Hurricane effect separately for each month. First I used January data in 2004 and

Table 2.3: Estimation on fertility effect by month: 2004 2006

	I	II	III	IV
Dependent variable: fertility rate*1000				
Treatment states:				
Primary	X	X	X	X
Secondary		X	X	
Control states:				
Adjacent	X		X	
Overall	X		X	
Year2006	0.2652***	0.1688***	0.2652***	0.1688***
	(5.69)	(5.45)	(5.34)	(5.51)
Treatment	-0.0725	0.109	-0.1937***	-0.0121
	(-1.18)	(1.01)	(-3.67)	(-0.16)
Jan*Y2006	0.0302	0.2311	0.00813	0.2103
	(0.15)	(0.6)	(0.05)	(0.77)
Feb*Y2006	-0.0689	-0.035	-0.0181	0.0262
	(-0.33)	(-0.09)	(-0.11)	(0.1)
Mar*Y2006	-0.1312	-0.1282	-0.0666	-0.0731
	(-0.64)	(-0.33)	(-0.4)	(-0.27)
Apr*Y2006	-0.2713	-0.3803	-0.2364	-0.3237
	(-1.32)	(-0.98)	(-1.43)	(-1.18)
May*Y2006	-0.0251	-0.245	-0.00058	-0.2113
	(-0.12)	(-0.63)	(0)	(-0.77)
Jun*Y2006	-0.0308	-0.0935	-0.0808	-0.1445
	(-0.15)	(-0.24)	(-0.49)	(-0.53)
Jul*Y2006	0.0289	0.0461	-0.0394	-0.0216
	(0.14)	(0.12)	(-0.24)	(-0.08)
Aug*Y2006	0.5882***	0.7201*	0.435***	0.5458**
	(2.85)	(1.86)	(2.63)	(1.99)
Sep*Y2006	0.3705*	0.6322	0.179	0.434
	(1.8)	(1.63)	(1.08)	(1.58)
Oct*Y2006	0.3829*	0.565	0.2582	0.4345
	(1.86)	(1.46)	(1.56)	(1.59)
Nov*Y2006	0.3496*	0.6747*	0.2241	0.5471**
	(1.7)	(1.74)	(1.35)	(2)
Dec*Y2006	0.1683	0.5612	0.0787	0.4748*
	(0.82)	(1.45)	(0.48)	(1.73)
agegroup	yes	yes	Yes	yes
state	yes	yes	Yes	yes
month	yes	yes	Yes	yes
R-square	0.9724	0.8694	0.9669	0.87171
N	1008	7056	1296	7344

2006 and estimated equation (2). Then I repeated the similar regression for the following 11 months. Table 4 reports the result using Primary Treatment States and Adjacent Control States. Thus for each regression shown in Table 4, theres no serial correlation problem as I used single pre-post periods; the coefficient on variable Treatment*year 2006 shows Katrina impact on fertility for every single month.

Table 2.4: Separate estimation on fertility effect: 2004-2006

	I	**II**	**III**	**IV**	**V**	**VI**
subsample	Jan	Feb	Mar	Apr	May	Jun
Year2006	0.1946	0.1047	0.1742	-0.0662	0.4065***	0.2694*
	(1.37)	(0.78)	(1.28)	(-0.48)	(2.94)	(1.77)
Treatment	-0.0974	-0.2003	-0.1386	-0.0855	-0.2397	-0.0502
	(-0.52)	(-1.14)	(-0.77)	(-0.47)	(-1.31)	(-0.25)
Treatment	0.0916	0.1313	-0.0237	-0.0665	0.0321	-0.052
***year2006**	(0.35)	(0.53)	(-0.09)	(-0.26)	(0.12)	(-0.18)
age group	yes	yes	yes	yes	yes	yes
cons	yes	yes	yes	yes	yes	yes
R square	0.9789	0.9775	0.9815	0.9781	0.9789	0.9769
N	84	84	84	84	84	84

continued

	VII	**VIII**	**IX**	**X**	**XI**	**XII**
subsample	Jul	Aug	Sep	Oct	Nov	Dec
year	0.2134	0.6976***	0.3697**	0.3443**	0.37**	0.1079
(0.75)	(1.48)	(3.66)	(2.11)	(2.03)	(2.22)	(0.75)
Treatment	-0.0838	-0.0142	0.0487	-0.0247	0.0697	-0.0512
	(-0.44)	(-0.06)	(0.21)	(-0.11)	(0.31)	(-0.27)
Treatment	0.0678	0.2877	0.2082	0.2966	0.1702	0.2411
***year2006**	(0.25)	(0.81)	(0.63)	(0.93)	(0.55)	(0.89)
age group	yes	yes	yes	yes	yes	yes
cons	yes	yes	yes	yes	yes	yes
R square	0.9807	0.9698	0.9739	0.9736	0.9737	0.9808
N	84	84	84	84	84	84

The coefficient γ for each month turned out to be insignificant using single-period estimation. Compared with result shown in Table 3, the treatment effect has significantly decreased from August to December. I conclude that in each separate month from 2004 to 2006, Katrina shows no significant impact on fertility for Hurricane affected states.

To further verify the problem caused by serial correlation, I estimated the following equations using data from June to December (assuming treatment effect will appear after

June) in 2004 and 2006 using two methods suggested by Bertrand et. al (2004):

$$fer_{ijtm} = \beta_0 + \beta_1 Y_t + \beta_2 T_i + \beta_3 M_m + \gamma(Y_t * T_i) + \beta_4 X_j + \epsilon \tag{2.3}$$

$$fer_{ijt} = \beta_0 + \beta_1 Y_t + \beta_2 T_i + \gamma(Y_t * T_i) + \beta_4 X_j + \epsilon \tag{2.4}$$

The first method is to use robust standard error clustered on state in multi-periods regression. Equation (3) used monthly data (multi-periods) and assumed the same treatment effect from June to December. The second method is to collapse multi-periods data into single pre-post periods. For equation (4) I collapsed data to single pre-post period (2004 and 2006) and used average monthly data. Column I of Table 5 shows result for equation (3) and Column II shows result for equation (4). Only regressions using primary treatment states and adjacent control states are report as other different specifications on treatment and control groups produce similar results. The coefficient of treatment effect is similar at 0.1742 and 0.1747 respectively, but the standard error in multi-period estimation is 0.123, much smaller than that in single-period estimation at 0.293, consistent with the theory that multi-period estimation may underestimate standard error even when it is clustered. This result again confirms that the overall Katrina effect on fertility is insignificant after correction.

Table 2.5: Estimation assuming same treatment effect from Jun to Dec

	I	**II**
year	0.3389***	0.339**
	(5.51)	(2.16)
Treatment	-0.0151	-0.0152
	(-0.19)	(-0.07)
year	0.1742	0.1747
*** Treatment**	(1.01)	(0.6)
age group	yes	yes
month	yes	no
cons	yes	yes
R square	0.9737	0.9776
N	588	84

This finding also cast doubt on previous paper by Rogers et. al (2005), which used multi-period birth data but did not apply any bias correction on standard error. The paper

may have overestimated the impact of Oklahoma City bombing.

2.4 County-level Analysis

To further examine the fertility effect of Katrina in all affected states, County-level analysis allows an analysis of hurricanes impact on fertility at different destruction level. As the NCHS collects county-level data only for counties with population of more than 100,000, it contains 60 counties in the four FEMA states Alabama, Florida, Louisiana and Mississippi.

Previous studies have used simple measures to proxy the local destruction like binary hurricane incidence or the level of hurricane advisories published in advance. However, the level of destruction may be various across different areas, and the aftermath caused by high wind blowing and consequential flood does not necessarily have linear linkage with the wind speed. A simple linear regression using dummy variables or wind speed may not be able to capture the nature of the catastrophe. Here I proposed two methods to proxy for hurricane destruction.

The first method to measure the level of local loss relies on the information revealed by the generosity of assistance offered by Federal Emergency Management Agency (FEMA). After Katrina, FEMA quickly evaluated the local destruction and designated many counties throughout the Gulf Coast as "disaster counties". This designation made these identified counties eligible to receive different level of disaster assistance from the federal government.

I group counties receiving FEMA assistance into 3 levels: Level 1, received only Public Assistance (PA) category A or B; level 2, received Individual Assistance (IA) and part of PA, or received full PA; level 3, received full IA and PA. I assume different assistance level corresponds to different degree of local destruction. I still use the fertility data from NCHS. Since NCHS only reports birth counts for big counties with population greater than 100,000, the data contains 60 counties in AL, FL, LA and MS.

I use the following equation to estimate the effect of hurricane:

$$\begin{aligned} fer_{itm} = \beta_0 + \beta_1 d1 + \beta_2 d2 + \beta_3 d3 + \gamma_1 d1 * year2006 + \gamma_2 d2 * year2006 \\ + \gamma_3 d3 * year2006 + \beta_4 year2006 + M_m + \epsilon \end{aligned} \tag{2.5}$$

Where fer_{itm} is fertility rate in county i in time t in month m. Its defined as monthly births per 1000 persons as population estimation on females by age is not available. d1 represents assistance level 1, d2 represents assistance level 2, d3 represents assistance level 3; base group is counties in AL, FL, LA and MS which received no FEMA assistance; year2006 is indicator that observation is from year 2006. I used interaction terms d1*year2006, d2*year2006, and d3*year2006 to capture the change of fertility caused by different level of hurricane loss. M_m is dummies for month. The analysis is restricted to observations only from June to December.

Using random effect and fixed effect model respectively, Table 6 Column 1 and 2 show the county level estimation result. The result indicates that areas received FEMA assistance have higher fertility rate in 2004 and 2006 but shows no different change after Hurricane Katrina, i.e. the interaction between treatment and after is not significant. This result is consistent with the result in previous section.

The second method involves a deeper understanding on hurricane damage. Strobl (2011) proposed an index to estimate the wind damage at the county level i using census tract level data. Both actual monetary loss and power dissipation in wind storms rises as the maximum wind speed increase and the validity of this index has been proved by previous studies (Emanuel, 2005; Strobl, 2011). I modified it to be applied to single hurricane strike:

$$Hurr_i = \sum_{j=1}^{n} V_{(i,j)}^{\lambda} w_j$$

where V is the maximum wind speed observed in census tract j in county i; w is the weight of census tract j in county i represented by some demographic characters. λ is a parameter that relates V to actual hurricane loss, which I set at 3.8 following previous geographic studies.

Hurricane speed V: Usually, historical data on hurricane reports only the path of the hurricane and wind speed of the eyes. Its difficult to determine the wind speed on a specific locale surrounding the eyes. A mathematical simulation could be a good alternative. An advanced model which simulates the full track of a hurricane storm was developed by Applied Research Associates during the period 1995-1997(Vickery, et al., 2000a, 2000b). Using

this approach, the dynamic of a hurricane beginning with its initiation over the ocean and ending with its final dissipation can be illustrated accurately. The model is evaluated as highly comprehensive by comparing the site-specific statistics of the key hurricane parameters of the simulated hurricane tracks with the statistics derived from the historical data (FEMA, 2007). This approach is implemented in a well-known geographic software HAZUS developed by FEMA to estimate the potential loss from the hurricane. One can generate the maximum wind speed (3-sec or 1-min) at census tract level using historical data of Atlantic hurricanes from 1900-2006.

Weight w_j: As the maximum wind speed generated from HAZUS is census tract specific, I use the ratio of population in census tract j to total population in county i as the weight.

After deriving $Hurr_i$, I run the following estimation:

$$fer_{itm} = \beta_0 + \beta_1 Hurr_i + \beta_2 Hurr_i * year2006 + \beta_3 * year2006 + M_m + \epsilon \tag{2.6}$$

Where the dependent variable fer_{itm} is the same as that in equation (5). $Hurr_i$ is the continuous variable measuring hurricane damage level in county i. The result is shown in Column 3 and Column 4 of table 6 using random effect and fixed effect model respectively. Since the coefficient β_2 is not significant, the null hypothesis that the level of hurricane damage does not have any impact on fertility rate cannot be rejected.

2.5 Robustness Check

There are certain concerns that estimation in previous sections might omit some important facts in reproduction. The effect of Hurricane Katrina was pervasive: several hospitals were flooded and experienced loss of functionality during Katrina. The recovery of hospital in severely flooded areas lasted for a long period, when many physicians and nurses relocated to less damaged areas (Berggren & Curiel, 2006). Its possible that pregnant women are more likely to move to unaffected areas to receive better pre- and post- natal care. In this case we may underestimate the increase of fertility rate caused by Hurricane Katrina. In contrast, another opinion is that pregnant women are less likely to move than average evacuees because the moving cost is higher for them; thus the increase in fertility rate might

Table 2.6: County-level Estimation

	I	II	III	IV
Disaster Level 1	0.2873***			
	(3.38)			
Disaster Level 2	0.1887***			
	(3.26)			
Disaster Level 3	0.2433***			
	(4.63)			
Hurr			.002***	
			(3.29)	
Disaster Level 1*Y2006	0.0161	0.0161		
	(0.68)	(0.68)		
Disaster Level 2*Y2006	0.0033	0.0033		
	(0.21)	(0.21)		
Disaster Level 3*Y2006	-0.0038	-0.0038		
	(-0.26)	(-0.26)		
Hurr*year2006			-0.00092	-0.00092
			(-0.62)	(-0.62)
year 2006	0.608***	.0608***	0.0645***	0.0645***
	(7.88)	(7.88)	(8.71)	(8.71)
month dummies	yes	yes	yes	yes
cons	yes	yes	yes	yes
N	840	840	840	840
counties	60	60	60	60
Random effect	Yes		Yes	
Fixed effect		Yes		Yes

be upward biased.

To check whether pregnant women followed different migration pattern from other women, I used 2006 American Community Survey (ACS) data to study females migration decision. The American Community Survey (ACS) is conducted every year and surveys approximate 3 million individual per year. The survey asked detailed questions on respondents household and individual characters. Also, it records the Public Use Microdata Areas (PUMAs) people lived 1 year ago and at present. I restricted analysis to females who lived in 4 FEMA states in 2005 and used multi-logit regression to check female migration decision.

Unfortunately, my birth data and FEMA area is by county but I have no information on counties of residence in ACS data. Therefore its necessary to link PUMA to county data to estimate the disaster level for each PUMA. I use weighted average of county disaster level (weight = county population in census year 2000) for PUMAs made up by several different counties. If a PUMA is out of FEMA areas, the disaster level equals to 0.

I category females into 4 groups: stay in same PUMA, moved to PUMA with same Hurricane disaster level; moved to PUMA with lower Hurricane disaster level; moved to PUMA with higher Hurricane disaster level. The explanatory variable of interest is whether the woman has children less than 1 year-old or not. Following a study on female migration decision by Enchautegui (1997), I used number of children aged from 1 to 5, age, age square, education level, employment, and family average income as controls.

Table 7 shows regression of females decision on migration. Comparing with staying in same PUMA, women who had children with less than 1 year-old are less likely to move to other PUMAs but the effect is not significant. The result indicates that moving propensity of pregnant women has little influence on my estimation.

2.6 Conclusion

This paper examines how Hurricane Katrina affects fertility rate in Gulf-coast states Alabama, Florida, Mississippi, and Louisiana. According to U.S. Vital Statistics Office, from 2005 to 2006 the average births per 1000 women increased respectively by 12.8% in Louisiana and 11.7% in Mississippi, significantly exceeding the U.S average fertility increase at 3.5%.

Table 2.7: Female decision on migration in AL, FL, LA, and MS in 2006

Independent Var	**Coefficient**	**Std. Err**	$P > \|z\|$
Base: Stay in same PUMA, omitted			
moved to PUMA with same Hurricane disaster level			
child less than 1 year-old	-0.159	0.135	0.237
# of children aged from 1 to 5	-0.126	0.036	0.001
age	0.219	0.019	0.000
age square	-0.004	0.000	0.000
education	0.109	0.009	0.000
employment	-0.317	0.038	0.000
average family income	0.000	0.000	0.000
cons	-5.781	0.244	0.000
moved to PUMA with lower Hurricane disaster level			
child less than 1 year-old	-0.108	0.122	0.375
# of children aged from 1 to 5	-0.267	0.070	0.000
age	0.176	0.034	0.000
age square	-0.004	0.001	0.000
education	0.164	0.016	0.000
employment	-0.446	0.068	0.000
average family income	0.000	0.000	0.000
cons	-6.657	0.437	0.000
moved to PUMA with higher Hurricane disaster level			
child less than 1 year-old	-0.126	0.099	0.201
# of children aged from 1 to 5	-0.047	0.051	0.351
age	0.170	0.026	0.000
age square	-0.003	0.000	0.000
education	0.105	0.013	0.000
employment	-0.590	0.054	0.000
average family income	0.000	0.000	0.000
cons	-5.746	0.339	0.000
N	57178		
LR	2153.78		

In this paper, a naive regression using difference-in-difference method and state-level aggregate birth count from Vital Statistics in 2005 and 2006 shows that there exists a significant increase in fertility in Hurricane Katrina affected states allowing for 9 months of gestation. However, I find that the naive estimation is problematic on both data and methodology. First, birth count data during the hurricane is abnormal due to evacuation; second, difference-in-difference method using multi-periods data will underestimate the standard error. The two issues are neglected by previous literature to some extent[1].

Using birth data in 2004 and 2006, I estimated the treatment effect using different model specifications. The result shows that the estimated standard errors are different in different regressions, which is consistent with the theoretical prediction by Bertrand et al. (2004). After using two correction methods proposed by Bertrand et al. (2004), I find that Hurricane Katrina had no significant impact on birth decision in those destroyed areas.

The county-level estimation is consistent with the conclusion reached by state-level data. I used two different methods to proxy for degree of destruction in different counties. The change of fertility in FEMA areas showed no difference from those unaffected areas; also, theres no relation between disaster level and fertility change.

[1]For example, Evans et. al (2004) which also used Vital Statistics did not question the data quality; Rogers et.al (2005) did not consider serial correlation problem

Chapter 3

How Responsive is the Supply of Chinese Education to Demographic Change: an Analysis Based on Hierarchy and Competition

3.1 Introduction

The educational system in China has expanded rapidly in recent decades. The enrollment rate in junior secondary school has risen from 77% in 1992 to about 100% in 2005 and subsequent years. The enrollment rates in senior secondary education and tertiary education have risen from 26.0% and 3.9% respectively in 1992 to 82.5% and 26.5% respectively in 2010. This expansion is likely to have been an important contributor to the rapid growth in Chinese GDP: the increase in human capital, including more universal access to education, create strong momentum for economic development in China and all around the world (eg. Barro and Lee, 1994; Li and Huang, 2009). In addition to boosting growth, the expansion in the supply of educated workers, viewed in isolation, would be expected to reduce the return to education. In fact, the return to education in China has continued to rise, with the premium for college graduates over high school graduates rising steadily from about 20 log points in 1988 to about 70 log points in 2008 (Han, Liu and Zhang, 2014). This is undoubtedly due to a concomitant rise in the demand for skilled workers caused by the transition from a planned to a market economy (Fleisher and Wang, 2005; Wu and Xie, 2003; Zhou, 2000).

The eagerness of students to acquire more education attests to rising demand. Its widely reported that the majority of third year students in middle school and high school are confined to school for more than ten hours every day to get higher scores in entrance exam to higher schools. In populated cities, parents pay as much as ten thousand dollars per year to procure access to elite public primary and secondary schools (Yang, 2011), for which

entrance is otherwise based on vicinity of residence or exam results. Even kindergarten graduates are reported to learn middle school materials to succeed in primary school admission exam. In 2003 the family expenditure on education accounted for 12.6% of the total family budget in Chinese cities (Yu and Suen, 2004).

It is therefore possible that the supply is still in shortage despite the expansion. To gain insight, I seek in this paper to understand the determinants of the supply of education in China and root causes of intense competition. No existing paper performs this economic and quantitative analysis, to my knowledge, although the intense competition has drawn sharp criticism from the general public and the academia. I consider the unique characteristics of the Chinese educational system: strict hierarchies and consequent competition among schools. Higher level schools, i.e. high school, and schools of higher quality, i.e. key-point schools, are at the top tier of the hierarchical system, able to enjoy better resources and serve more capable students. I begin with a theoretical model, in which I suggest that when schools and education authorities make joint decisions on enrollment, they face a tradeoff between quantity, since both individuals and government's long-term objective of economic development demands high enrollment, and high quality, since school officials need to provide high quality relative to other schools to vie for more government appropriation and attract more talented students. My model predicts that the optimal levels of enrollment and quality for revenue-constrained schools are determined by three major arguments: demand, resources available, and the school's specific weighting of enrollment (w) versus quality (1-w). Generally schools on top tier of the hierarchy set a higher weight on quality. I then provide empirical analysis examining the responsiveness of the supply of education, especially to changes in cohort size, and also test the model by estimating the impact on enrollment and quality of relative weights schools put on enrollment versus school quality.

The benchmark empirical analysis is Chinas secondary schools enrollment elasticity with respect to cohort size, underscoring the mechanism that constrained supply creates crowding-out effect in educational institutions[1]. There are only two papers studying the cohort size effect in mainland China: Coxhead, Shen, and Yao (2014) found that baby

[1]The cohort size effect on education is complex due to varying assumptions and background conditions.

boomers born after 1961 benefited from large-scale school construction in 1970s and had higher primary school enrollment rate. Using individual level data, Ma (2014) found that one born in smaller cohorts compared to future cohorts enjoy more years of education. Compared to their reduced-form analysis, my analysis is based on schools optimal choice on enrollment quota and I find that in general a 10% increase in cohort size reduces middle school enrollment rate by less than 1% and high school enrollment rate by 3-4%. Meanwhile, a 10% increase in cohort size reduces teacher/student ratio in primary school, middle school and high school by 3.9%, 3.7% and 2.5% respectively.

To further evaluate my theory, I consider variations in weight on quantity (w) versus quality ($1 - w$) across time and provinces for the same type of school. Though w cannot be estimated directly without assuming utility function form or having individual school's data, I find variables that may affect w or serve as an indicator of w and examine whether the effect of changes in those variables is consistent with the effect of changes in w as is predicted by the model. In middle school level, since the implementation of *Compulsory Education Law* is essentially a process for the authorities to put higher weight on enrollment than quality, I estimate its impact on enrollment and quality and find that, as predicted, it lowered fat cohorts loss in terms of enrollment opportunities but exacerbated their disadvantage in terms of quality. In high school level I use the number of high schools per county as an indictor for w. I find that one additional high school increases enrollment elasticity with respect to cohort size by 0.1 while reducing quality elasticity with respect to cohort size by .04. I also find that officially designated key-point high schools, who have high criterion on quality, have much lower enrollment elasticities than ordinary high schools. Those findings

---

One strand of literature focuses on the long term interaction between demographic change, choice of education level, and wage structure, underscoring the imperfect substitution between young and old workers (Connelly, 1986; Easterlin, 1978; Easterlin, Wachter, and Wachter, 1978; Falaris and Peters, 1992; Fertig, Schmidt, Sinning, 2009; Flinn, Freeman, 1979; Stapleton and Young, 1988; Wachter and Wascher, 1984; Welch, 1979; etc). The Chinese scenario is arguably more related to another strand of literature focusing on how an abrupt increase in cohort size worsens students opportunities due to limited resources. For developed economies, as most of them have realized universal secondary education, researchers focus on cohort crowding effect on college attainment (Bound and Turner, 2006; Card and Lemieux, 2001). In lower level schools, though resource available may affect the grade retention, the effect is minor (Babcock, Bedard, and Schulte, 2012) or even positive (Reiling, 2016). For developing countries, Saavedra (2012) documented policies in several countries, suggesting that restricted access to public schools are the major reason of cohort crowding. He found that in Columbia a 10% increase in cohort size decreases the completion rate of high school by 3%, that of college by 4% and average schooling years by 1%. Huang and Xie (2014) found similar results using samples in Taiwan.

add evidence on literature about schooling quality (Barro and Lee, 2001; Card and Krueger, 1996; Hanushek, 1992; Harmon and Walker, 2000; Jensen, 2010). Though a set of papers discussed about quality and quantity trade-off, most are from the view of individual's decision making process (see e.g. Guo, Yi and Zhang (2016) for a review). This paper offers a specimen where the government, as a social planner, makes tradeoffs between quantity and quality and shows how educational institutions followed the government policy to realize their specific objectives.

Overall, my results suggest that Chinese schools responsiveness to increased demand for education are consistent with the objectives set by different schools at different time. In such a typical hierarchical system where resource distribution to different schools is extremely uneven, higher weight is put on quality $(1 - w)$ for more favored schools, i.e. key-point schools and higher level of schools. On the supply side, this is the reason for rationing access to schooling above middle school; on the demand side, this is the reason for intense competition among students. The selective and competitive nature revealed in this paper challenges the practice to directly applying traditional Mincers method to measure return to education in China, since China has a totally different education system compared with countries where access to school, at least before tertiary stage, is not constrained by supply[2]. Though a large body of literature has calculated return to education in China, only in a recent paper by Li, Liu and Zhang (2012) the impact of selective educational system was mentioned. They found that compared to other countries in China a very small proportion of return to education (2.7% out of 8.4%) is a result of accumulated human capital and the rest is caused by omitted ability, suggesting non-tertiary education in China serves to pick up the most capable ones rather than teaching students skills used in workplaces. Given the characteristics of Chinas education system, I also develop a signaling model about the selecting process of higher level schools and prove that this process leads to a welfare loss for students.

The remaining parts of this paper proceed in the following way: Section 2 provides

[2]The traditional Mincer function is based on the assumption that the present discounted value of going to school is equal to that of not going to school. Even many papers consider selection and sorting biases, they still assume that individuals have greatest sway in determination of their optimal schooling years.

background information, including a panorama on fertility change in China in sixty years and a brief instruction to characteristics of education system in China. Section 3 presents a model showing how schools make a tradeoff between enrollment and quality. Section 4 use census data to show enrollment elasticity for different schools in different historical periods. Section 5 use richer institutional data in recent decades and used additional variables to proxy parameters required when schools make decisions. It shows how enrollment and quality in different levels of schools respond to a set of changes. Section 6 is a case study to show the hierarchies within high school system and their effects on admission decisions. Section 7 analyzes how the existence of limited enrollment and hierarchy influences students' learning behavior and further their total welfare. Section 8 provides the conclusion.

3.2 Background

3.2.1 A brief timeline about contemporary China: Fertility change and educational attainment

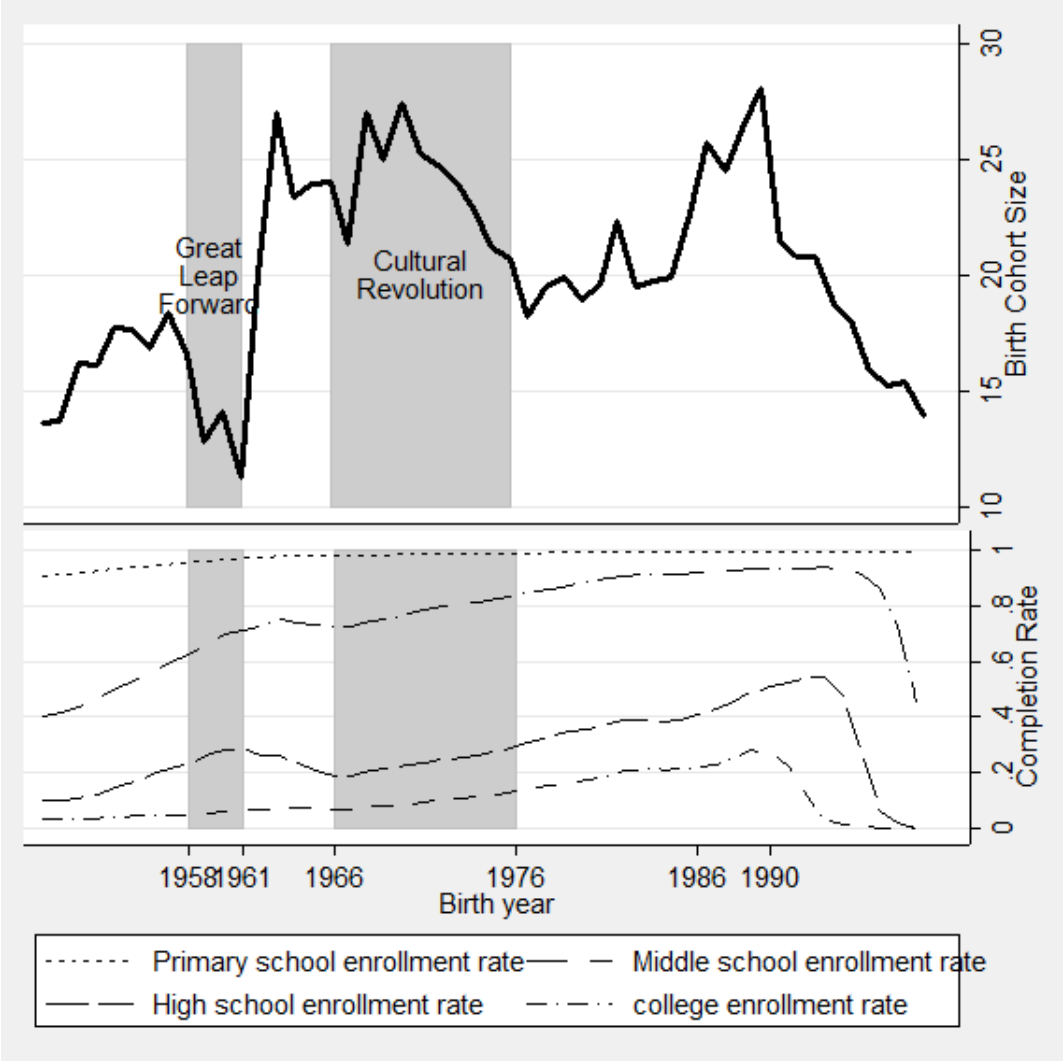
Chinese history in second half of the last century is marked by drastic changes in mainly two respects: transition of socioeconomic regimes and fluctuation in population. I will enumerate historical landmarks and events that have potential impacts on the whole economy and labor market outcome as well as trend of birth population.

Figure 1 lays out the change in birth cohort as well as some key events. Focusing on demographic cycles, the history went through the following periods:

1. *Steady growth of population: 1949-1957.* After the finding of Peoples Republic of China, there was a period of reconstruction when the economy was recovering from decades of war and accompanied expansion of birth population. During this period, China was experiencing high fertility rate of about 5 to 6 children per woman (Ping, 2000).

2. *Sharp declines caused by calamity: 1958-1961.* In 1957, the so-called Great Leap Forward occurred. Due to a series of policy errors, the agricultural production plummeted and the economy was stagnant. A great drought exacerbated the calamity, leading to a sharp drop in fertility and excess deaths. The total fertility rate fell from 6.2 in 1957 to 3.3 in 1961 (Ping, 2000) and the birth population dropped from 18 in 1957 million to 11

Figure 3.1: birth cohort population and percentage of enrollment



million in 1961.

3. *First baby boom caused by high reproduction rate in communist movement: 1962-early 1970s.* The economy began to recover from 1962. The birth cohort in 1962 (25.6 m) doubled than that in 1961. From then on, nearly one decade of baby boom followed, creating a huge cohort of .24 billion people born from 1962 to 1970, even after the destructive Cultural Revolution has started in 1966.

4. *The implantation of birth control policy: early 1970s to mid 1980s.* Though most functions of the administration system was suspended during Cultural Revolution (1966-1976), a notable policy of this period is that starting from early 1970 a national family planning program was introduced to harness the population growth out of control, including later marriage and birth, longer spacing, and fewer births(Chen, 1984). The total fertility rate decreased to 2.7 children per woman in early 1980s whereas the drop in birth population was not rapid enough due to large population base of women in reproductive age.

5. *Second tide of baby boom caused by first boom: 1986-1990.* The economic reform after 1978 is characterized by strict birth control. The one-child policy was formally introduced in 1979 as a Basic State Policy. Despite the strict policy, the population is the next ten years were in stagnation due to large population base. The 1986-1990 cohort was regarded by researchers as the second generation baby boomers as the first generation baby boomers entered reproductive age.

6.*Low fertility period: 1991-present.* With strict birth control policy, the birth population continued to decline after 1991. The increasing educational attainment of women also contributed to the trend (). The total fertility rate was only 1.5 children per woman in 2000 and 1.7 children per woman in 2010 (World bank, 2015).

Figure 1 clearly reveals a dramatic change in birth population - two tides of baby booms and one period with relatively low fertility rate in between the two baby booms. Although only one major alteration on fertility policy occurred during the past 60 years, the course of history exerted profound influence on fertility behaviors, creating some experimental variation on cohort size.

3.2.2 Education in China

As is well known, to receive education in China means a great deal of pressure, efforts, and competition. Its widely reported that the majority of third year students in middle school and high school are confined to school for more than ten hours every day to get higher scores in promotion exams. The competition starts at very young age: in populated cities, there are anecdotal stories that kindergarten graduates learn middle school materials to succeed in primary school admission exam. A series of studies on education and sociology documented Chinese educational system and reached unanimous agreement that Chinese students were faced with intense competition at varying degrees, from primary school to college (Lin and Chen, 1995; Fleisher, Li, Li, and Wang, 2004; Han and Yang, 2001). The question then arises: why students in China show such a high level of competitiveness compared to peers in other countries? A simple answer could be that education supply in China is scarce so students have to make extra efforts for limited slots. This is true in early periods. However, with the expansion of schooling system and increasing government appropriation concurring with the growth of Chinese economy in recent decades, theres no sign that the frenzy for education began to quell. The key point here is the competition for high quality rather than simply seeking for admission. Several specific characters of Chinese education system can help us understand the rationality in behaviors of all parties, including parents, schools, and local authorities.

a. *One-time exam oriented system.* After the foundation of Peoples Republic of China in 1949, government devoted to eliminate illiteracy and providing more education to the public. Due to constrained budget and lack of eligible teachers, all students were free to enter primary school but have to pass a set of strict exams to be enrolled into middle school, high school and college in sequence. The idea underpinning such exam mechanism is that education is not seen as a mean of edifying the broad masses of people, but rather of selecting the most capable to lead the majority and of enhancing mobility between different social hierarchies. This situation began to alter in 1987, when the Compulsory Education Law was promulgated. The Compulsory Education Law stipulates that all school-age children must receive nine-year compulsory education, including six years in primary school and three

years in middle school. As the policy phased in, government adopted a series of policies to remove barriers to enter schools and the compulsory education was completely free of tuition charge since 2008. After the completion of middle school, however, education is no longer universal and students who wish to continue their study must pass a unified entrance exam to be enrolled by senior secondary schools, including general high schools and vocational schools. The admission is based on students ranking in the exam as well as the slots of senior secondary schools. After three years studying in senior high schools, graduates pursuing more education are required to take the unified National College Entrance Exam (NCEE) to be admitted by tertiary educational institutions. Therefore, the most important, if not the only, criteria in Chinese system is high score in unified one-time entrance exam[3].

b. *Organization.* Its understandable that in the densely-populated developing country, the public expenditure on education is far from sufficient to accommodate the high demand for education. Some developing countries have a higher proportion of private schools as a supplement to public education system (Chen, 2005; Glewwe and Kremer, 2006; James, 1993). China, however, did not have any private schools until the early 1980s and currently sees a very low proportion of private school still.

The public schooling system is dominated by local education bureaus. Its responsibility includes allocation of appropriation, designation of headmasters, and the organization of teacher hiring. Different schools, according to their importance and universality, are in the charge of different government tiers.[4] The prevalent practice in rural area is that primary, middle, and high schools are funded and administered by village, township and county government while the best schools are mostly directly administered by county government; in city areas, schools are administered by district government with the exception that best schools are directly administered by prefecture bureaus (Tsang, 1996). The number of different schools is consistent with the institutional arrangement: as is shown in Panel a

[3]There's widely discussion that transcript during school time should be included as a part of admission criteria, in practice the enrollment decision is solely determined by scores in one-time entrance exam in most areas.

[4]China has a strict hierarchy that central government is in charge of provincial administrations and then prefecture administration. A prefecture has city areas composed of several districts and rural areas composed of several counties as a counterpart. Under administration of county government are township and then village.

of Table 1, though the number of schools fluctuate across year, the number of schools in different levels has totally different order of magnitude. On average a county has hundreds of primary schools. They are widely distributed in the big rural area and administered by lower education bureaus in order to provide easy access and edify the mass. There are fewer middle school and even fewer high school in each county[5] and the status of school becomes higher as theyre administered by higher education bureaus and their headmasters and teachers have higher ranking in the administrative ladder. Typically, a county with population of half million has less than 10 high schools. Many students choose to board at school due to far distances to higher level of schools, especially in rural area, resulting in a great rate of boarding students. In 2010, 21.97% primary school students are boarding students and 43.67% middle school students are boarding students[6][7]. Statistics for high school is absent but the figure is expected to be higher and most high schools offer boarding choice in order to accommodate students from across the county and save more commuting time for laborious coursework.

Teachers educational attainment also varied across different level schools. Panel b of Table 1 indicates that high school teachers in general received more education. In 1990, 45.5% high school teachers earned bachelors degree and more 90% have degrees above college. At this year 46.6% middle school teachers and 5.7% primary school teachers received college education. Though teachers educational attainment increased quickly in the following 20 years, there still exists substantial gaps in education received between high school, middle school and primary school teachers. A large proportion (21.7%) of primary school teachers have only received upper secondary education or even less, while the proportion of high school teachers received more than bachelors degree becomes 94.8% in 2010.[8]

---

[5]As a comparison, for U.S. from 1980 to 2012 the ratio of number of primary school to secondary school is 2.42 to 2.66 and theyre not administered by different tiers of government

[6]Source: 2010 National Statistics Report on Education Development

[7]A set of government documents state that building more boarding school and provide better boarding condition are long term objectives of education development in order to provide easier access for rural students. For example, see *Outline of Mid-and-Long Term National Education Reforms and Development Strategy: 2010-2020*

[8]As a comparison, in 2011-2012 the proportion of highest degree earned by U.S. elementary and secondary school teachers are respectively 3.8%, 39.9%, 47.7%, 7.6%, and 1.1% for less than bachelors degree, bachelors degree, masters degree, education specialist, and doctors degree. Among teachers for 9th-12th graders, the figures are 4.9%, 38.2%, 47.9%, 6.8%, and 2.1%. There exists no systematic difference between high school

Table 3.1: Primary and secondary school: number and teachers education level

Year	1980	1990	2000	2010
a. Number of schools				
High school	31300	15678	14564	14058
Middle school	87077	71953	62704	54823
Primary school	917316	766072	553622	257410
b. Teachers' highest degree earned				
High school				
Bachelor's or above		45.5%	68.4%	94.8%
some college		45.6%	30.2%	5.1%
upper secondary or below		8.9%	1.3%	0.1%
Middle school				
Bachelor's or above		6.9%	14.2%	64.1%
some college		39.7%	72.9%	34.6%
upper secondary or below		53.5%	12.9%	1.3%
Primary school				
Bachelor's or above		0.3%	1.0%	23.7%
some college		5.4%	19.0%	54.6%
upper secondary or below		94.4%	80.0%	21.7%

Source: *Chinese Education Statistics Yearbook*

c. *Hierarchies within same level of schools.* The so-called Key-point school system stemmed from 1950s, immediately after the foundation of PRC. As China was extremely poor then, the budget-constrained governments had different attitudes to lower level of education and higher level of education: lower-level education should be universal to enlighten the mass while higher-level education should be elitist to select the most capable to serve as specialist in favor of technology development[9]. As qualified teachers and related resources are scarce, budget-constrained government officially categorized all schools into key-point (Zhongdian) schools and ordinary schools. This policy was reinforced in reform period since 1980[10]. Resources from government are intensively directed to key-point public schools. The key-point schools are able to receive more funds and enjoy a series of privileges like better

teachers and other teachers in elementary and secondary schools. Source: *Digest of education statistics 2014*. National Center for Education Statistics, 2015.

[9]In a report in 1956, then president Liu Shaoqi said: "Universal education is not too urgent now: the question now is still higher education and the need for specialists" (Chandra, 1987)

[10]A series of policies were promulgated by Minister of Education around 1980 and related documents include: *A trial scheme on running a number of key-point primary and secondary schools* (1978); *Resolution on running key-point schools by stage* (1980); *Opinions on improving quality of middle shool and high school* (1983)

teachers and better infrastructures.[11] After the *Compulsory Education Law* was enacted in 1987, officially key-point primary and middle school regime was gradually removed in the following decade. However, the cumulated advantage continued and students still flock to former key-point schools which thereby could set high criteria on admission or charge extra money. In high school level, the key-point school system continued and thrived. A large body of literature in Chinese throw into sharp relief the differences in resource acquired by key-point and ordinary schools (see Yang (2006) for a review).

There're also hierarchies within key-point schools and ordinary schools. Official tournament among all schools in the same level are organized by local education bureaus, which is possibly the root of intense competition in Chinese schools. It has long been questioned why Chinese government officials have incentives to work hard in absence of democratic voting system. Zhou(2004, 2007) proposed tournament mechanism as an institutional arrangement to explain this. In China's top-down hierarchical government institution, government officials' interest lies in their promotion in the hierarchy, which is tightly associated with upper-level governments' evaluation on their performance. Zhou (2007) suggests that such tournament systems have the following characters: the system should be centralized so upper-level administration can determine the interest of subordinates; there exists objective criteria so that all participating parties consent to the result of competition; the efforts (or decision) of subordinates can affect the result of competition; it's essentially a zero-sum game if the total reward to competition is fixed. Those characters were indeed shared by education system.

In China upper-level education bureaus organize tournament among subordinate bureaus, and lower-level education bureaus organize tournament among subordinate schools in same levels following similar logics. Every year after the high school and college entrance exams, middle schools and high schools are ranked respectively according to their

[11] In an interview with education authorities (Mauger, 1983), the officials regarded the key-point schools and the examination system for entry to higher education as regrettable necessities: "The institution of key-point schools is a last choice when we have no other alternative. In the future, with the provision of enough qualified teachers and enough equipment we can adopt an equal approach towards all children. By weighing up the advantages and disadvantages of enrolment examinations we have come to the conclusion that we must stick to the criterion of quality first. So the educational administration and the educationists must have a sober understanding of this situation.

students' performance in the exam. Same level schools with better academic performance, namely promotion rate to higher level of schools, could enjoy more government appropriation and more promotion opportunities for the headmasters. This simple assessment may affect school officials' behaviors in two aspects: first, they may want to limit enrollment partly to guarantee enough resource per student and partly to select students with higher endowment. Second, they may pass the objectives to teachers and lead similar tournament among teachers, who directly exert pressure on students to achieve better academic performance[12]. It's common practice that teachers' wage and bonus are associated with the performance of their students.

The tournament among education bureaus is more complicated as measures are needed to expand enrollments as well as improving quality. If upper-level government demands expansion of educational opportunity and uses this as a criterion, the subordinate bureau will also pass this objective to schools. The common strategy to such policy objective is that key-point schools keep limiting enrollment in order to keep quality and ordinary schools choose to expand enrollment.

The tournament mechanism has been widely cited to explain the regional economic growth and provision of public goods in China. Zhou, Zou and Chen (2013) found evidence that there exists such tournament among local governments on education expenditures but how this kind of tournament affects enrollment and schooling quality remains elusive.

Although different schools offer different quality, the nature of public school determines that they officially charge the same tuition[13]. In this situation, school entrance exam serves as the only mean to pick out the most intellectual children who would enjoy the privilege of receiving more and better education. Students compete to enter higher level of schools, and, at the same time, they compete to enter a better school among all level of schools, as a student in a better middle school is more likely to be admitted by a better high school, then a better college. Note that the relationship between students and higher schools are

[12]This may also explain why in China colleges do not exert much pressure on students. Without a simple measurement of promotion rate, the evaluation on college performance is much more complicated and students' effort has relatively small impact.

[13]On facing with constrained budget, government encourages public schools to raise funding by all means. It is common practice that schools admit high scored students charging low tuition and low scored students charging high tuition. See later section about funding for more discussion.

two-way selection: students would compete to enter a school with better prestige, more government investment and those schools, in return, would pick better students to perform better in government evaluation and receive more funds. A natural result is that the initial advantage would accumulate and further widen the performance gap. Once the hierarchies among different schools are formed, students gain full information about their quality as well as expected effort required to enter those schools; also schools are very concerned about quality by picking out the best students and limiting enrollment to some degree.

Due to this mechanism, theres huge difference on academic performance, i.e. promotion rate to higher schools, among schools. This could be reflected by a comparison between key-point schools and ordinary schools. For example, in a key-point school in Guangzhou city 72% of high school students went to college, against the rate of about 3.8% for overall high school promotion rate in the province. In Sichuan only 22% of all high school graduates are from key-point schools, but they make up 90% of college entrants. (Chandra, 1987; Mauger, 1983; Rosen, 1985)

d. *Funding.* Each year the birth population in China is from 15 to 28 million in past half century. Considering the relatively high primary and middle school enrollment rate, the resource needed by education system is vast. On the other side, however, education expenditure as a fraction of GNP and total national budget is unduly low. In 2000, government expenditures on primary schools and secondary schools are respectively .79% and .89% of GNP (computed using data from Chinese Education Expenditure Yearbook, 2000), while at the same time the figures are 1% and 1.1% for low-income countries, 1.8% and 1.4% for middle-income countries, and 1.4% and 1.9% for high-income countries (Glewwe and Kremer, 2006).

The funding of schools mainly comes from the following channels: appropriation from central government, tax revenue collected by local government, and schools revenue generated from teaching activities including tuition, miscellaneous fee, and school selection fee. After cultural revolution, since the limited government budget can hardly reconcile with high demand for education both on individuals side and nations side to promote economic development, from 1980s the government encourage schools to generate more funding by charging a certain amount of tuitions. Despite this, the central government also set a long

term objective to offer free nine-year education in 1987 *Compulsory Education Law*. With two decades of economic growth, rural students receiving nine-year compulsory education are exempt for all expense in 2006 and urban students are exempt in 2008. Table 2 shows the education cost and sources in different levels of schools. The data is generated from aggregated expenditure data for 31 provinces/municipalities. Table 2 shows that cost per

Table 3.2: Average Expenditure and Sources of Funding

	1998		2004		2010	
	Mean	SD	Mean	SD	Mean	SD
Primary school						
Expenditure per student	616.6	198.8	1342.7	542.9	3221.2	1051.9
% government appropriation	72.8%	6.1%	81.9%	4.5%	95.1%	2.6%
% student payment	14.2%	4.1%	12.2%	3.3%	3.1%	2.0%
Middle school						
Expenditure per student	1043.0	362.8	1609.4	729.7	4238.4	1441.1
% government appropriation	69.3%	5.4%	73.8%	4.8%	92.5%	2.8%
% student payment	16.6%	3.8%	18.7%	3.6%	5.3%	2.1%
High school						
Expenditure per student	2561.1	831.7	3228.8	1221.3	5394.6	1860.8
% government appropriation	55.7%	6.6%	49.9%	5.7%	65.5%	6.2%
% student payment	27.1%	5.5%	39.7%	6.3%	31.1%	5.5%

Note: Weighted mean and standard deviation across 31 provinces/municipalities. Source: *China Educational Finance Statistical Yearbook*

student and percentage paid by students themselves increase as a student receive higher level of education. Even when government appropriation alone is considered, the funding tips in favor of high school student disproportionately. As mentioned before, the reason is schools in higher level are administered and funded by government in higher level and therefore receive more funding. Also, high schools charge more tuition and have additional sources of revenue: the so called *school selection fee* [14]. Every year after the unified high school entrance examination, a high school release two threshold on exam scores. A student scored more than the higher threshold would enter the publicly-funded trajectory and pay only a low tuition. A student scored less than the higher threshold but more than the lower threshold would enter the private-funded trajectory by paying tuition plus expensive school

[14]Some key-point primary and middle schools also charge school selection fee to accommodate students who are eager to go to those schools. Excess demand is key to this surcharge.

selection fee, which is always several times of tuition. As lowly ranked schools have lower criteria, a student qualified for private funding in a highly ranked school could either pay the expensive school selection fee or go to a school with lower rank paying only tuition. As high school is in general scarce resource, all high schools, ranked high or low, have considerable revenue on school selection fee.[15]

Per student expenditure continued to rise with time as a result of more government dedication. In primary and middle school level, the proportion of expenditure funded by students themselves decreased from 12.4% and 15% in 1998 to 2.2% and 4% in 2010 respectively, confirming the government promises to turn compulsory education free of charge. In high school level, however, as high school is not included in compulsory education students keep paying a large amount of money, a majority of it being school selection fee.

The brief summary on primary and secondary education reveals that Chinese education system is highly hierarchical twofold. It's hierarchical across different levels of schools. A higher level of school is more elite. High schools have more stringent admission criteria, enjoy more qualified teachers, and receive more appropriation than middle schools than primary schools. It is also hierarchical in the same level of schools. A key-point school has higher enrollment thresholds and enjoy better resources than an ordinary school. We may conclude that a school at top tier of hierarchy values more about quality and relatively a school at lower tier of hierarchy put more weights on enrollment.

3.3 Econometric Model of School Behavior

Consider the decision making process of schools. This is essentially not a standard private market as cost paid by students is only a small proportion of total cost and a school doesn't aim at profit-maximizing. A school (under the direction of local authority) aims at accommodating students demand for education while keeping a certain level of education quality Q. Education quality is closely related to resources allocated to each student. As a non-profit organization, a public school naturally faces a binding budget constraint that cost to all students to attain a certain level of quality should be equal to tuition

[15]Due to paucity of data the exact number or scale of school selection fee is not attainable, so Section 6 offers a case study about school selection fee.

plus government appropriation ($gG(n)$) (Hansmann, 1981)[16]. $G(n)$ is the component of government appropriation that varies with n. Its proper to assume that as the number of students n increases, government funds increases at decreasing rate, i.e. $G(n) > 0, G''(n) < 0$ [17]. g is a coefficient on G which represents the variation independent of n[18]. Denote tuition for each student as T and the budget constraint is $c(Q, n) - T * n - gG(n) = 0$. I assume $C_Q > 0, C_n > 0, and C_{QQ} > 0, C_{nn} > 0$ and $C_{Qn} = 0$ (separability).

Assume a simple linear objective function for a school:

$$U(Q, n) = (1 - w)Q + w\alpha(N) * n$$

A school aims at achieving high quality as well as enrolling more students (with the constraint that enrollment is smaller than or equal to applicants, $n \leq N$). w represents the weight on enrollment and different schools may value quality and enrollment at different weight. $\alpha(N)$ is a coefficient function of N. As we assume that schooling system adjust admission slots to accommodate students' demand, $\alpha'(N) > 0$ to represent increasing marginal utility of enrollment as N increases.

Then we have Lagrangian:

$$L = (1 - w)Q + w\alpha(N) * n - \lambda[C(Q, n) - T * n - gG(n)]$$

To maximize utility, the first order condition is that at optimal (Q^*, n^*):

$$\frac{\partial L}{\partial Q} = 1 - w - \lambda C_Q = 0$$

[16]Though some literature documented the trend of privatization in education (Mok, 1997), referring to the opening of private schools after later 1980s, its not a major concern in our paper for the purpose of selectivity, as private schools are always associated with low quality of students enrolled and generally regarded as a supplement to public schools.

[17]The assumption that government expenditure is relatively inflexible holds in Chinese context. Unlike the flexible schooling system in U.S. where the funding of public schools is generated from property tax, the source of primary and secondary school funding is government appropriation in province and municipality level in China

[18]Government expenditure on education depends on financial means of local government as well as cohort size. Therefore g could captures, for example, increase in per-student expenditure with time and inter-regional differences in resource per student

$$\frac{\partial L}{\partial n} = w\alpha(N) - \lambda[C_n - T - gG'] = 0$$

$$\frac{\partial L}{\partial \lambda} = C(Q, n) - T * n - gG(n) = 0$$

Therefore we have

$$(1 - w)/[w\alpha(N)] = C_Q/[C_n - T - gG'(n)]$$

This means that the relative marginal utility of quality and marginal utility of quantity should be equal to the ratio of their marginal cost. It holds that $C_n - T - gG'(n) > 0$ at optimum since the increase in funding cannot catch up with increase in total cost with quality unchanged. The Hessian is:

$$\begin{bmatrix} -\lambda C_{QQ} & -\lambda C_{Qn} & -C_Q \\ -\lambda C_{Qn} & -\lambda(C_{nn} - gG'') & -(C_n - T - gG') \\ C_Q & C_n - T - gG' & 0 \end{bmatrix} \tag{3.1}$$

Since $\lambda > 0, C_{Qn} = 0$, and $C_n - T - gG'(n) > 0$ at (Q^*, n^*), the Hessian is negative definite at (Q^*, n^*). There exists local maximum.

First, consider how w affects equilibrium $n*$ and $Q*$:

$$\frac{\partial n}{\partial w}|_{n=n^*} = \alpha'(N)C_Q\{w^2[\alpha'(N)]^2C_{QQ} + (1 - w)^2(C_{nn} - gG'')\}^{-1} > 0$$

$$\frac{\partial Q}{\partial w}|_{Q=Q^*} = -\frac{w}{1 - w}[\alpha'(N)]^2C_Q\{w^2[\alpha'(N)]^2C_{QQ} + (1 - w)^2(C_{nn} - gG'')\}^{-1} > 0$$

This shows that as w increases, i.e. the school attach more importance to enrollment rather than quality, more students will be enrolled and quality will be sacrificed.

Consider two extreme scenarios $w = 0$ or $w = 1$.

When $w = 1$, it means the school aims at maximizing enrollment. Therefore a corner solution where $n = N$ is the school's optimal choice.

When $w = 0$, the school's objective is to maximize quality. According to the budget constraint, quality is maximized when $C_n - T - gG' = 0$[19], i.e. when the marginal cost to

[19]The FOC is $\frac{\partial Q}{\partial n} = -\frac{C_n - T - gG'}{C_Q} = 0$, and the second derivative is $\frac{\partial^2 Q}{\partial n^2} = -\frac{C_{nn} - G''}{C_Q} < 0$, suggesting a local maximum

extra student is equal to increase in fundings.

As demand for education (N) changes, the comparative statistic is:

$$\frac{\partial n}{\partial N}|_{n=n^*} = \frac{w}{1-w}\alpha'(N)C_Q\{(\frac{w}{1-w})^2[\alpha'(N)]^2C_{QQ} + C_{nn} - gG''\}^{-1} > 0 \tag{3.2}$$

$$\frac{\partial Q}{\partial N}|_{Q=Q^*} = -(\frac{w}{1-w})^2\{\alpha'(N)]^2C_Q[(\frac{w}{1-w})^2[\alpha'(N)]^2C_{QQ} + C_{nn} - gG''\}^{-1} < 0 \tag{3.3}$$

Since the elasticity is $\eta = \frac{\partial n/n}{\partial N/N}$, the range of w determines elasticity of a school. The key-point schools are quality maximizer, meaning $w \to 0$, therefore $\eta \to 0$ [20] and the quality change goes to zero too. Ordinary schools are quantity maximizers; this leads to the corner solution that $n = N$. In this scenario, the real enrollment is determined by demand side. If the proportion of students who want to go to a higher school is fixed, the real enrollment elasticity will be 1. If a schooling system is a mixture of key-point school (quality maximizer) and ordinary school (quantity maximizer), the overall elasticity will range between 0 and 1.

Note that g is also a determinant for quality and quantity:

$$\frac{\partial n}{\partial g}|_{n=n^*} = [G' + \frac{w}{1-w}\frac{\alpha C_{QQ}G}{C_Q}\{(\frac{w}{1-w})^2\alpha^2 + C_{nn} - G''\}^{-1} > 0$$

meaning that as government expentiture increases, there will be more students enrolled. The sign of $\frac{\partial Q}{\partial g}|_{Q=Q^*}$ is ambiguous.

Since the parameters and functional form are not observable, I decide to use two reduced form equation to generalize the determinants of schools' choice, with some implications on derivatives generated from above model.

$$n = n(N, w, g)$$

[20] When $w \to 0$, $\frac{w}{1-w} \to 0$, so

$$\eta \approx o(\frac{w}{1-w})\frac{N}{n}\alpha'(N)C_Q\{(o[\frac{w}{1-w})^2][\alpha'(N)]^2C_{QQ}+C_{nn}-G''\}^{-1} \approx o(\frac{w}{1-w})\frac{N}{n}\alpha'(N)C_Q(C_{nn}-G'')^{-1} \approx 0$$

$$Q = Q(N, w, g)$$

The above model implies that $n_N > 0$, $n_w > 0$, $n_g > 0$; $Q_N < 0$, $Q_w < 0$, and the sign of Q_g is ambiguous.

For the purpose of estimating elasticity of enrollment and quality with respect to cohort size when w and g is absent, the empirical strategy is to run the following regression using aggregate data:

$$lnEnrollment^s_{cp} = \beta^s * lnCohSize_{cp} + \delta^s_p + \rho^s_c + u^s_{cp} \tag{3.4}$$

$$ln(Tch/stu)^s_{cp} = \gamma^s * lnCohSize_{cp} + \delta^s_p + \rho^s_c + u^s_{cp} \tag{3.5}$$

where $s = \{$Middle school, High school, College$\}$. The dependent variable $Enrollment_{cp}$ is the total enrollment of cohort c in province p for middle school, high school and college. The dependent variable Tch/stu is a proxy for quality, teacher to student ratio. $CohSize_{cp}$ is a proxy for N, representing cohort size of a birth cohort. A year dummy is added to control for year fixed effect and province dummy is added to control for province fixed effect. Note that the estimated coefficient β_1 represents the elasticity of school enrollment with respect to cohort size.

$$\beta_1 = \frac{\partial lnEnrollment}{\partial lnCohortSize} = \frac{\partial Enrollment/Enrollment}{\partial CohortSize/CohortSize}$$

The null hypothesis is $\beta_1 = 1$, meaning percentage change in enrollment is equal to percentage change in cohort size and the schooling system can flexibly adjust to demographic change. Though w didn't appear in the regression, it does affect elasticity. In the hierarchical system of China, a school in higher tier of hierarchies would set a higher weight on quality $(1 - w)$ rather than quantity(w). Therefore, high schools are supposed to have lowest enrollment and quality elasticity with respect to cohort size; in the same schooling level, key point schools are supposed to have lower enrollment and quality elasticity than ordinary schools.

The fixed effect estimation using equation (4) and (5) is unbiased only when the province and cohort dummies can fully capture the variation in w and g across time and provinces, which is not necessarily true. So I need add proxies for w and g if data permits. Though

the effect of w and g is entangled with N rather than a simple addition, I can still run a linear regression as an approximation to check the marginal effect.

$$lnEnrollment^s_{cp} = \beta^s_1 * lnCohSize_{cp} + \beta^s_2 * w_{cp} + \beta^s_3 * g_{cp} + \delta^s_p + \rho^s_c + u^s_{cp} \tag{3.6}$$

$$ln(Tch/stu)^s_{cp} = \gamma^s_1 * lnCohSize_{cp} + \gamma^s_2 * w_{cp} + \gamma^s_3 * g_{cp} + \delta^s_p + \rho^s_c + u^s_{cp} \tag{3.7}$$

3.4 Empirical Results Using Census Data

Since data on government appropriation is absent in early periods, I use equation (4) to check the impact of cohort size on the educational system by seeing the final educational outcome shown in census data. The bottom panel in figure 1 shows the big picture of population and educational attainment by birth cohort using 2010 census data. Though there exists an upward trend in educational attainment, the frequent fluctuation in birth population offers an opportunity to observe the effect of cohort size disentangled from time trend.

The data I use is 2010 census data assembly at province level. Though micro level data is unavailable, the dataset provides the counts of population by educational attainment, age, and province. The questionnaire asks respondents about the highest level of school she or he has ever been enrolled[21] and tabulate the number of enrollment in each birth cohort for each province. The province-level data is favorable due to the following reason: First, theres inequality across province and government appropriation on basic education was mainly determined by local government in a long period. According to Tsang (1996), since the early 1980s, the funding of basic education in China has tipped away from a centralized system and towards a decentralized system with a diversified revenue base, mainly relying on local fiscal revenue and revenue collected by schools themselves. The advantage of such a system is that schools have discretion on planning admissions to adapt to varying population. The disadvantages, on the other hand, are that local government

[21]The dataset provides two counts on educational attainment: the highest level of school one has been enrolled and the highest level of school one has completed. As this study focuses on supply side decision, i.e. whether schools enrolled students or not, rather than students dropout decision, I use enrollment counts by education level instead of completion

may be reluctant to invest in education and that there will be huge inequality caused by unbalanced development level across the nation. Using this dataset, table 3 shows the level of inequality in educational attainment across 31 provinces/municipalities in selected years.

Table 3.3: Inequality in educational attainment across provinces

Birth year	Mean	SD	Max	Min	N
1950	6.92	1.33	9.76	2.69	31
1960	8.67	1.28	11.27	3.65	31
1970	9.04	1.06	11.41	4.90	31
1980	9.98	1.26	13.04	5.47	31

Source: China 2010 population census data assembly, 2011

Second, the labor market mobility across province is relatively low. Since the aggregate data reported in census is based on the population who has been living in one place for more than half a year, rather than by place of registration (hukou) or by birth place, there exists a potential source of bias in estimation that ones location at census time may be different from her or his location when going to school. This bias is possibly relieved by Chinas institutional restriction on migration. The 2010 census data shows that 92.0% of the population still lives in the province same with their birth province. Generally, the proportion is smaller for more developed provinces (From lowest to higher: Zhejiang, 76%; Guangdong, 77%; Xinjiang, 83% .) and larger for more populous and less developed provinces (Sichuan: 98%, Hunan: 98.5%, Henan: 98.8%).

I run the regression for different periods in contemporary Chinese history and schooling levels separately. I would assume that during a short period the demand for education by different cohorts should be homogeneous, i.e., theres no endogenous change in each individuals demand for education caused by variation in cohort size in a short period and the proportion of students who want to go to a certain level of school should be close across years in a short period[22]. Therefore, if the elasticity is not equal to 1, it simply reflects

[22]There are two main channels through which the demand may change with cohort size. First, boomers and busters may face different conditions, mostly importantly the return to education. This is why literature sees education as endogenous variable in analysis about cohort size's impact on wage (for detailed discussion, see Connelly, 1986). In China case, however, as boom and bust did not last long and the wage system is relatively fixed, the effect of demographic shock may not pass to labor market so as to change educational demand. Also, relatively short time spans are chosen in order to avoid this kind of endogeneity that may exist in long time span. Second, even in short time span adjacent cohorts may still have different individual's

adjustment on supply side. For division of "short periods", I follow the normal division by historians (the historical events and the basis of division are as introduced in Section 2). As fertility changes are not concurrent with social changes, there are enough variations in cohort size in each period.

Table 4 shows the results. Using equation (1), column I, II, and III represent determinants of enrollment in middle school, high school and college respectively. The time periods in different panels indicate the time when an individual entered specific school, assuming students enter middle school, high school and college at age 12, 15, and 18 respectively. For example, the cohort born in 1960 is supposed to be included in the samples who go to middle school from 1966 to 1976 (Panel c) and also in the samples who go to college after 1978 (Panel d). A province fixed effect model is applied and robust standard error clustered on province is chosen to control for serial correlation within a province.

In the period immediately after foundation of PRC (1950-1957), as shown in Panel a, the effect of cohort size on middle school enrollment rate is .911, insignificantly smaller than 1. This means when cohort size increases by 10%, enrollment increases by 9.1%, showing that the enrollment in middle school is relatively elastic. The elasticities of high school and college are close, meaning a 10% increase in cohort size will lead to a 5% increase in high school and college.

Panel b shows the elasticity of admission during turmoil before Cultural Revolution. The elasticity for high school increased. As we can see from the magnitude of coefficient, middle school is more elastic than high school and college. This means large cohort is more disadvantaged in admission of higher level school.

Panel c shows the elasticity of admission during Cultural Revolution. Its widely known that during Cultural Revolution (1966-1976) all educational institutions were under attack. In the peak time of the movement from 1966 to 1968, normal teaching activities in all kinds of schools were interrupted (Unger, 1982). For basic education, it was common practice for primary school age students to stay at home and for junior and senior high

---

demand since large cohort size dilutes the education resource per student (Bound & Turner, 2006) and lower their willingness to obtain more education. There's no good way to distinguish this effect but if we assume demand is excessive this kind of effect may be minor.

Table 3.4: Enrollment Elasticity of Cohort Size in Different Periods

	I	**II**	**III**	
	Middle school	High school	college	
Dependent variable: ln(total Enrollment)				
a. 1950-1957: after foundation of PRC				
LnCohortSize	0.911	0.545***	0.585***	N
	(0.0597)	(0.0858)	(0.0808)	248
R-sq	0.985	0.977	0.988	
b. 1958-1965: turmoil before Cultural Revolution				
LnCohortSize	0.856**	0.778***	0.536***	N
	(0.0725)	(0.0850)	(0.0845)	248
R-sq	0.981	0.896	0.785	
c. 1966-1976: Cultural Revolution				
LnCohortSize	0.939	0.428***	0.181***	
	(0.0382)	(0.0824)	(0.130)	341
R-sq	0.974	0.950	0.935	
d. 1978-2004: Post-reform period				
LnCohortSize	0.913***	0.691***	0.612***	N
	(0.0302)	(0.0938)	(0.0810)	837
R-sq	0.937	0.906	0.957	

Note: * p<10% ** p<5% ***p<1%. The significance level on elasticity estimation is against the hypothesis that $\beta = 1$. Covariates include birth cohort size, province dummy and cohort dummy. Fixed-effect on province is applied. Within group R-squares are reported and robust standard errors clustered on province are in parentheses. Source: China 2010 population census data assembly, 2011

school and university students to meet each day at school to engage in political activities (Meng and Gregory, 2002). After schools' reopen in 1968, the teaching activity was still in disorder under the environment of general social turmoil. However, figure 1 proved that at least the general admission for those susceptible cohorts was not impacted: there isnt a sudden dip in middle school enrollment rate for 1954-1956 birth cohorts (who are supposed to attend middle school from 1966 to 1968, the most destructive period in Cultural Revolution). Analogously, there isnt a sudden dip in the high school enrollment rate for 1951-1953 birth cohorts. Table 4 shows that during Cultural Revolution middle school enrollment is inelastic, while its still close to 1. For high school, however, the elasticity is only 0.430, meaning a 10% increase in cohort size will lead to a 4.3% increase in the enrollments of high school, which is equivalent to a 5.7% decrease in enrollment rate. This indicates a decrease in the nations ability to adjust enrollment policy to accommodate varying population. The elasticity of college enrollment is even lower at .272.

Panel d shows the elasticity of admission during reform period. The effect at middle school level is similar to previous period. The coefficient of high school is 0.665. It means a 10% increase in birth population will lead to a 3.3% decline in high school enrollment rate, which is approximately 1 percentage points decrease in high school enrollment rate as the average enrollment rate is 30.6% from the data. Similarly, a 10% increase in birth population will lead to a 4.2% decline in college enrollment rate, which is approximately .6 percentage point decrease in college enrollment rate as the average enrollment rate is 13% during this period.

3.5 Empirical Results Using Institutional Data

3.5.1 cohort size and enrollment

Compared to census data shown in Section 3, institutional data on true enrollment and graduates can better capture the decision made by supply side. As richer data is only available in recent decades, I use aggregate data collected from *Chinese Educational Statistics Yearbook (1987-2013)* and *Chinese Educational Expenditure Statistics Yearbook (1987-2013)* to further explore enrollment and quality elasticity in reform period. Still the data is aggregate

on province level. More detailed data, i.e prefecture or county level data is not available across the country.

I first run regression using equation (4) but different data. The dependent variable $Enrollment_{cp}$ is the enrollment number of middle schools or high schools in province p in year t. If the dependent variable is middle school enrollment, then the independent variable $CohortSize_{cp}$ represents primary school graduates in that year to proxy demand for higher level of education. If the dependent variable is high school enrollment, then the independent variable $CohortSize_{tp}$ represents middle school graduates in that year. The estimated coefficient β represents the elasticity of school enrollment with respect to the number of graduates from lower level of schools. Also, its worth noting that middle school graduates may go to secondary technical school as well as high school, the former being regarded as a same level as high school and registered in census indiscriminately. I also calculated the elasticity of technical school enrollment with respect to middle school graduates as Im unable to differentiate between middle school graduates willing to go to two types of higher schools.

Its necessary to check the elasticity of college enrollment, yet data on college enrollments by province is absent. Though theres statistic about enrollments of colleges located in one province, its inappropriate to use such data to estimate enrollment rate as a large part of students go to college in other provinces. Therefore, for college admission I just use enrollment in institutional data as side evidence.

In Table 5 Panel a shows the elasticity of middle school, high school, secondary technical school and college during the period from 1987 to 2013. The enrollment elasticity decreases as level of school increases since a lower level of school attach more weight on enrollment, consistent with the model prediction. Also, the comparison between high school and secondary technological school shows that ordinary high school is less elastic, as high schools are in the academic track and schools care more about quality as promotion rate to universities is an important criterion to evaluate their performance.

Panel b shows the comparison between institutional data and census data in a comparable period. As the census was launched in 2010 and some children may delay entering school, the enrollment statistics for late cohorts who are supposed to enter school at around

Table 3.5: Institutional data: elasticity of enrollments with respect to graduates †

	I	II	III	IV	V
	middle school	ordinary high school	secondary technology school	overall senior secondary school	college
a. Elasticity: 1987-2013					
	0.957	0.655***	0.794*	0.695***	0.430***
	(0.0645)	(0.0527)	(0.113)	(0.058)	(0.0816)
N	702	702	656	656	700
R-sq	0.935	0.970	0.871	0.961	0.989
b. Elasticity: 1987-2004					
Institutional data	0.892*	0.542***	0.629**	0.572***	0.305***
	(0.0654)	(0.0662)	(0.163)	(0.0931)	(0.125)
N	468	468	465	465	466
R-sq	0.936	0.950	0.772	0.933	0.983
Census data	0.919*			0.770**	0.878*
	(0.0369)			(0.0823)	(0.0764)
N	468			468	468
R-sq	0.975			0.940	0.973

Note: Independent variable is enrollment and covariates include graduates of lower level school, province dummy and cohort dummy. Elasticity represents coefficient of graduates. * p<10% ** p<5% ***p<1%. The significance level on elasticity estimation is against the hypothesis that $\beta = 1$. Within group R-squares are reported and robust standard errors clustered on province are in parentheses. Source: China 2010 population census data assembly, 2011; Chinese Educational Statistics Yearbook (1987-2013)

2010 may be biased downward in census data, especially for college data. Therefore I choose institutional data up to year 2004 to generate a reliable comparison between census and institutional data[23].

The institutional data shows much smaller elasticities for both middle school and high school enrollments than census data. This might be explained by self-adjustment by students themselves. It's common practice that if a middle school graduate failed to enter a target high school, he may choose to sit another year in middle school and retake the entrance exam next year. Therefore, the adverse enrollment opportunity faced by disadvantaged big cohorts could be mitigated at the cost of their extra effort and time. Thus the elasticity in terms of overall admission shown in census data is greater than the elasticity in terms of enrollment in a specific year shown in institutional data. Though comprehensive data on scale of retaking class and entrance exam is absent, there's plenty of side evidence. For example, in 1999 for all students who registered for college entrance exam, 31% of total have completed high school in previous years and determined to take the exam again. The number rise to 36% for rural students. In 1993, 10% of students who registered for high school entrance exam in Shenyang city have completed middle school in previous years.

3.5.2 schooling quality and cohort size

Another concern is how expansion of enrollment opportunities changed the quality of schooling. I use the teacher- student ratio to proxy schooling quality. The data I used is obtained from *Chinese Educational Statistics Yearbook (1987-2013)*, which contains statistic on primary and secondary education students and teachers by province. The year books have statistic on full-time teachers and substitute teachers respectively. Briefly speaking, substitute teachers are analogous to part-time lecturer, receiving lower wage and compensation than full time teachers.

Figure 2 shows the trend of teacher/student ratio and the number of students in primary school by province. Its quite obvious that the two curves move in opposite direction. As

[23]For census data, I assume all students attend middle school at 12 and high school at 15. Therefore for middle school the sample includes cohorts born from 1975-1992, for high school the sample includes cohorts born from 1972 to 1989. Certain provinces absent in institutional data are also deleted from census sample.

the student number increased, the teacher/student ratio dropped immediately, suggesting an inelastic supply of teachers that doesn't respond to variation in the number of students.

In table 6, column I, II, III and IV shows the elasticity of full-time teachers number with respect to student number. The estimates show that as number of students increases teacher/student ratio decreases and the elasticity is smaller for lower-level schools. For primary school, a 10% increase in student number will cause a 6.1% increase in teacher number. For middle school, this figure is 6. 3% increase. The elasticity in high school is higher at 0.749, showing a smaller sacrifice in quality as quantity increases. In column V and VI, I include substitute teachers and use the number of all teachers as dependent variable. As the yearbook only have statistic on substitute teachers in middle school and high school together, I can only do regression of ordinary secondary school teachers on ordinary secondary school students. The comparison between I and V shows that the inclusion of substitutes teachers increased the elasticity in primary school by 10% (from 0.614-0.678). Column IV and VI show the inclusion of substitute teachers slightly mitigated the negative impact on large cohort in secondary school though it's not significant.

The following regression is considered:

$$lnTeacher_{tp} = \beta_1 lnStudent_{tp} + dumm.province + dummy.year + u_{cp} \tag{3.8}$$

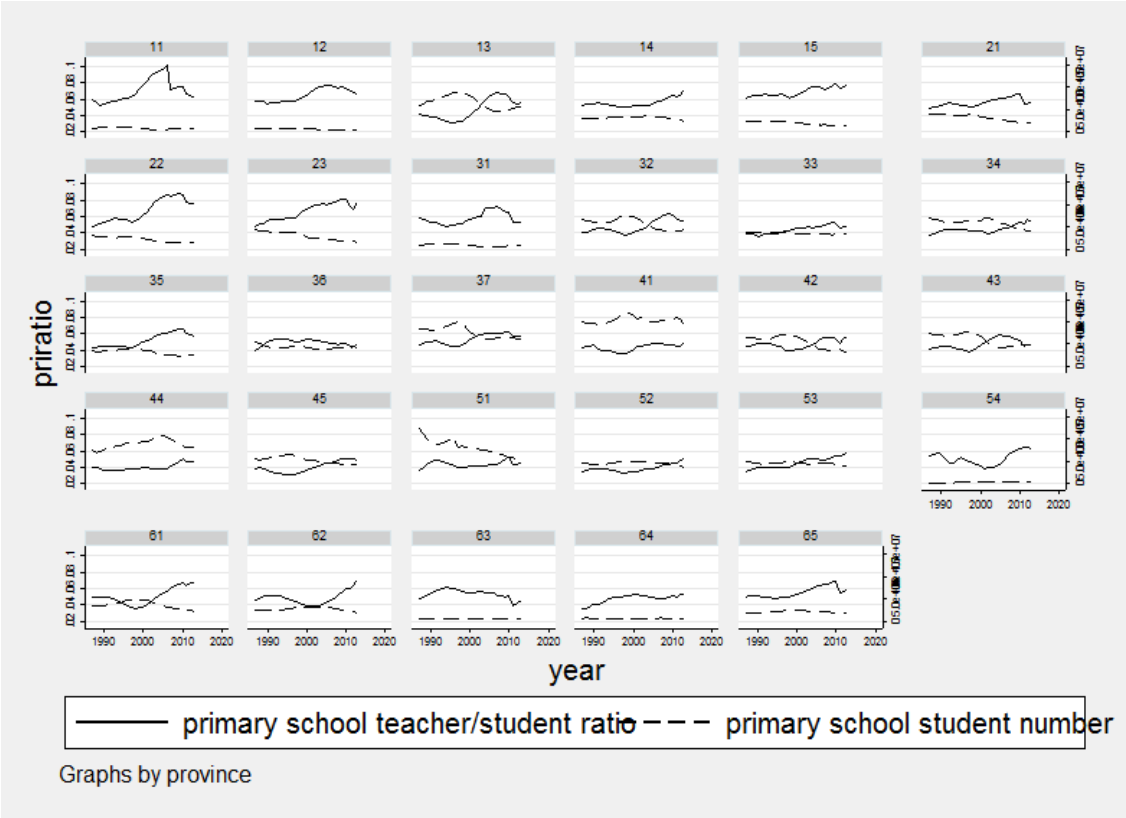


Figure 3.2: Primary school: student number and teacher/student ratio by province

Table 3.6: Elasticity of teachers number with respect to students number

	I	II	III	IV	V	VI
	primary school	middle school	high school	ordinary secondary school	primary school	ordinary secondary school
Dependent variable: *Ln(Teacher)*						
Independent variable:						
Ln(Student)	0.614***	0.634***	0.749 ***	0.678***	0.671***	0.690 ***
	(0.0879)	(0.0490)	(0.0748)	(0.0636)	(0.0630)	(0.0612)
Full-time	Yes	Yes	Yes	Yes	Yes	Yes
Substitute	-	-	-	-	Yes	Yes
Year Dummy	Yes	Yes	Yes	Yes	Yes	Yes
Province Dummy	Yes	Yes	Yes	Yes	Yes	Yes
R-sq	0.559	0.904	0.975	0.941	0.777	0.948

Note: * p<10% ** p<5% ***p<1%. The significance level on elasticity estimation is against the hypothesis that $\beta = 1$. Fixed-effect on province is applied. Within group R-squares are reported and robust standard errors clustered on province are in parentheses. Source: Chinese Educational Statistics Yearbook (1987-2013)

Theres major difference in the patterns of enrollment elasticity and quality elasticity. The enrollment elasticities decrease with schooling level while quality elasticities increase with schooling level. With limited resources, public school system is always facing a trade-off between quantity and quality. In primary and middle school stage, the government made effort to expand enrollment to accommodate more students and the enrollment elasticity is close to 1. However, the quality is, to some degree, sacrificed. As for high school stage, since government put more weight on quality, enrollment elasticity is lower but quality is more steady across cohorts.

3.5.3 Factors affecting quantity-quality tradeoff

As richer data on education expenditure is available after 1995, I could use it to check the determinants of enrollments, including demand for education, government expenditure, and competition level. In the theoretical session, the assumption is that the funding of education cannot keep pace as enrollment increases. I used education expenditure data from 1995 to 2011 to check if this assumption holds.

Growth on Educational Expenditure

As is stated in Section 2, the funding of schools mainly comes from the two channels: first, government appropriation, including appropriation from fiscal revenue and ad-hoc education-related tax[24] revenue collected by local government; second, schools revenue generated from teaching activities including tuition, miscellaneous fee, and school selection fee. Thus I estimated the following function to examine how fundings change as number of students changes.

$$lnFunding_{tp} = \beta_1 * lnStudent_{tp} + \delta_p + \rho_c + u_{cp} \tag{3.9}$$

The independent variable $lnStudent_{tp}$ represents log of students registered in year t in province p. The dependent variable $Funding$ here includes four different variables:

[24]The major source of such tax funding is called *Education Supplymentary Tax*, which is now 3% of the sum of value-added tax, business tax and sales tax. The education supplymentary tax is directly paid by local business and entirely used on education.

- *TotRev*: total revenue of schools
- *GovApp*: government appropriation, including fiscal appropriation and education-related tax revenue
- *SchInc*: income generated by school itself, including tuition, miscellaneous fee, and school selection fee
- *Tuition*: tuition directly paid by students

TotRev is mainly composed of *GovApp* and *SchInc*, and *SchInc* includes *Tuition*.

Table 3.7: Education Funding on Student Number

	I	II	III	IV
Dependent variable:	*LnTotRev*	*LnGovApp*	*LnSchInc*	*LnTuition*
Independent variable:	lnStudent			
a. High School				
lnStudent	0.716*	0.536***	1.015	1.098
	(0.143)	(0.129)	(0.174)	(0.218)
R-sq	0.975	0.979	0.952	0.939
b. Middle School				
lnStudent	0.471**	0.428***	0.623*	0.437**
	(0.195)	(0.154)	(0.209)	(0.245)
R-sq	0.939	0.976	0.573	0.511
c. Primary School				
lnStudent	0.283***	0.235***	0.928	1.141
	(0.145)	(0.115)	(0.315)	(0.369)
R-sq	0.966	0.978	0.506	0.529

Note: * p<10% ** p<5% ***p<1%. The significance level on elasticity estimation is against the hypothesis that $\beta = 1$. Fixed effect model is estimated and robust standard errors clustered on province are reported in parentheses. Source: Chinese Educational Statistics Yearbook (1987-2013); China Educational Finance Statistical Yearbook (1987-2013)

Table 7 shows how school funding changes as number of students changes for high school, middle school and primary school. I use province fixed effect and year dummy is added to capture year difference. Panel a, b, and c show the elasticities of school revenue with respect to number of students in high school, middle school and primary school respectively. In general, when number of student increase by 10 percent, high schools' total revenue increases

by 7.2 percent, middle school increase by 4.7 percent and primary school increase by 2.8%. Obviously when number of students increases, the total revenue generated by schools are not enough to produce a persistent quality across cohort. Different levels of schools show different responsiveness. High schools' total revenue is most responsive, followed by middle schools and primary schools. The differences are determined by two causes: first, as a major source of funding, government appropriation received by high school is more elastic at .54 and that received by primary school is least elastic at .28. This result is consistent with fact presented in section 2 that higher level of schools enjoys more funding and better resources, suggesting that government set more premiums on quality of education in higher levels.

Second, the composition of funding varies across different school. As is shown in Table 2, the proportion of funding paid by students themselves is about 3-14% in middle school and 5-17% in middle school and around 30% in high school. Different sources of funding show quite different patterns: the elasticity for government appropriation is the least at .45 while elasticities for income generated from students are slightly greater than 1. The comparison indicates that public funding is quite insensitive to change in number of students and students in large cohorts paid more to compensate for the lack of public funding.

Since government expenditure increases at a decreasing rate as enrollment increases, basic assumption in theoretical model holds. The estimation shown in table 2 can also help identify g. As is shown in theoretical model, the government expenditure is a product of G, government expenditure that depends on n, and g, coefficient on G that represents different scale of government expenditure across province and time. As equation (9) disintegrates $Funding$ into a part determined by n and error terms, essentially

$$lnG_{tp} = \beta_1 * Student_{tp}$$

and

$$lng_{tp} = \delta_p + \rho_c + u_{cp}$$

Then g could be estimated using the estimated coefficient shown in Column II of table 7. As $n = n(N, w, g)$ and $Q = n(N, w, g)$ the question now boils down to find out variables that capture variations in w.

Supply of Middle School Education

As has been stated in model part, the objective set by government and constrained by government budget plays a key role in determination of enrollment and schooling quality. To prove this, the snag is to find w - weight on enrollment - for each province.

Theres no policy that has more fundamental impact on middle school enrollment in China than the *Compulsory Education Law* enacted in 1986. Its the first time that China enacted a law to specify education-related policies. The major stipulations in this law include: first, a nine-year compulsory education would be gradually implemented in the whole country; second, all children at age 6 are required to be enrolled into compulsory education[25]; third, the state will appropriate fund for the improvement of compulsory education and guarantee the growth rate of educational expenditure would be higher than that of total fiscal revenue[26] .

Following 1986 compulsory education law by central government, all province governments formulated blueprints and passed acts on how to promote compulsory education, specifying objectives and concrete measures to carry out the law. Educational administrations try to raise more money to fund schools with the help of local government and urge middle schools in their charge to expand enrollment. This process is also followed by reallocation of educational resources, teachers and government appropriations included, so that the gap between key schools and ordinary schools diminished to accommodate students indistinguishably. Finally in areas that satisfy certain conditions to apply nine-year education[27], all pupils are supposed to be exempt from middle school entrance exam and allocated to schools in vicinity of their residency.

As all students are allowed to enter middle school according to vicinity of residence, the policy brings about a major change in middle schools enrollment decision: schools put all

---

[25] In practice for less developed areas children can start primary school at 7.

[26] The law prescribes that compulsory education is free of charge. However, this article wasnt applied until the Amendment to Compulsory Education, passed in 2006, included tuitions of compulsory education into government budget. See China state council, 2008.

[27] The decree issued by central government is that local authority should determine to cancel the exam when the goal of middle school prevalence (enrollment rate higher than 95%) is accomplished. In practice local governments have the discretion.

weight on quantity and admit as many students as there are once the entrance exam is cancelled. The progression of compulsory education is essentially a general process that in more regions schools begin to set $w = 1$. Therefore, I suggest a linear mapping between w and the progression of universal nine-year compulsory education in a specific province: the closer to the achievement year, the larger the overall w in that province[28]. I use the variable years from the achievement of universal compulsory education to proxy the progression[29]. For example, if one province achieved universal compulsory education in 2000, then for year 1995 I assign a value of $1995 - 2000 = -5$ for the variable Progression and for years after 2000 the variable is just equal to 0.

Though the exact year in which a province achieved universal compulsory education is available[30], it may suffer the bias of reverse causality that arouses when a shock in enrollment caused a faster progression rather than the other way around. Thus I use the ex-ante expectation on the year in which a province realizes universal compulsory education, which is set as a government goal in related government or administrative documents. Most provinces announced the year in which compulsory education should be accomplished in provincial acts passed soon after national *Compulsory Education Law*. For the few provinces without a specific year in provincial acts, I used related information recorded in education administrations documents. As the four municipalities are developed areas and have already realized universal 9 years education early, I excluded them from my sample and leave only data in 27 provinces.

---

[28]Note that the progression represents a continuous change in w rather than binary, because a province can claim to achieve universal education only when all its counties achieve universal education and there exist disparities among counties within the same province. In practice a county has the discretion to remove exam system when it has full capacity to enroll all school-age students. Therefore the progression affects the proportions of counties having achieved universal middle school education in a specific province

[29] Previous literatures (Fang et.al, 2012, Song, 2012) focus on the implementation date of related acts enacted by provincial government. However, the implementation date may not be a good proxy of policy progression. For example, the act in Shandong province was effective on Sep 9th, 1986, 3 days prior to Sep 12th, 1989, the effective date of the act in Jiangsu province. This subtle difference is not a source of variation when estimating the policy effect, as a new school year always starts on September 1st. However, the schedules set up in the two acts are quite different. Jiangsu act specified that nine-year education should be universal before 1995 whereas Shandong act specified that nine-year education should be extended to hinterlands before 2000. The effect also turned out different; Jiangsus enrollment rate increased by 5.5 percentage point in 5 years whereas that of Shandong increased by 13 percentage point.

[30]Regular practice is that officials from national and provincial education board evaluate the yearly progress of Nine-year compulsory education penetration county by county and approve on whether one area has accomplished the universal education.

As the initial situation varied across the whole country, the goal set by provincial governments differed a lot. Provinces with good initial condition would set the target year in between 1992 to 1995; about half of provinces, which have medium condition, mostly set the target year at 2000 (or by the end of the (last) century); backward provinces, mostly in west area, would leave the details fuzzy (at least for progression of middle school) in initial acts but later (at around 2000) set it between 2005 and 2010. The following figure shows the actual progression by provinces setting different goals. Enrollment rate is calculated as the number of middle school enrollees over the number of primary school graduates.

Figure 3 shows that provinces followed their schedules to a large extent. The high level provinces reached an enrollment rate greater than 95% in 1995. The medium level provinces, regardless of their variation initially, has enrollment rate greater than 90% in 2000. Backward provinces also reached enrollment rate greater than 95% in 2005[31].

I regress school enrollment (n) and teacher/student ratio (Q) on N, w, and g using equation (6) and (7). Table 8 shows the regression result using different settings. The first Column shows that in general the enrollment elasticity is .957 in the period from 1987 to 2013. Column II added progression of compulsory education as an independent variable and Column III added coefficient for government appropriation and the sample shrank to 1995-2013 due to paucity of data. Column IV to VII use schooling quality as the dependent variable. It's clear that as the number of middle school graduate increases, enrollment elasticity is about 1, indicating no sacrifice of quantity, at the cost of significant quality decrease. When primary school graduates increase by 10%, middle school enrollment increases by 10% while teacher/student ratio decrease by 4%. The coefficient on the variable *progression* shows that as the government tries to extend compulsory education, more students were enrolled and the quality was sacrificed. With each additional year of progression, enrollment increases by 1.6% and teacher/student ratio decreases by 1.2%. Note that the *progression* may have two separate effects: a change in w as well as an increase in g since government promises to raise more funding when phasing in compulsory education. Therefore I add lng in Column III and VI to disintegrate the effect of *progression*. If g can

[31]Data in Qinghai seems to show abnormality as it has high enrollment rate as early as late 1980s. However this is because primary school in Qinghai was not universal, not indicative of high education level.

Figure 3.3: Middle School Enrollment Rate by Different Provinces

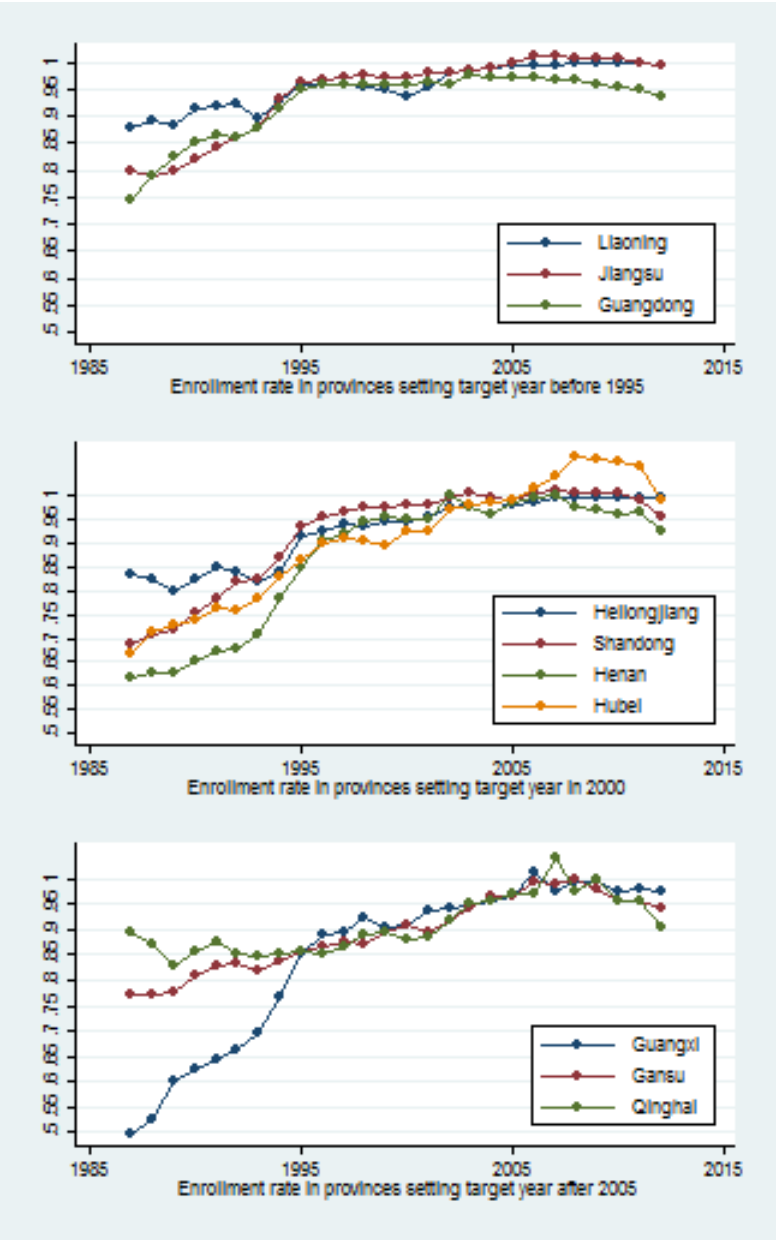


Table 3.8: Determinants of Enrollment and Quality in Middle School Education

	I	II	III	IV	V	VI
Dependent variable:	LnEnrol	LnEnrol	LnEnrol	Ln(Tch/Stu)	Ln(Tch/Stu)	Ln(Tch/Stu)
Independent variable:						
LnGraduate	0.957	0.961	1.028	-0.305***	-0.307***	-0.404***
	(0.0643)	(0.0455)	(0.0345)	(0.0542)	(0.0380)	(0.0568)
Progression		0.0167***	0.00933**		-0.0116**	-0.00765
		(0.00336)	(0.00419)		(0.00474)	(0.00572)
Lng			0.0688			0.012
			(0.0451)			(0.052)
Year dummy	yes	yes	yes	yes	yes	yes
province dummy	yes	yes	yes	yes	yes	yes
N	702	702	468	702	702	468
R-sq	0.935	0.973	0.975	0.828	0.842	0.899

Note: * p<10% ** p<5% ***p<1%. The significance level on elasticity estimation is against the hypothesis that $\beta = 1$. Fixed effect model is estimated and robust standard errors clustered on province are reported in parentheses. g represents government funding independent of number of students. Source: Chinese Educational Statistics Yearbook (1987-2013); China Educational Finance Statistical Yearbook (1987-2013)

fully capture the variation of government funding, regardless of the source of variation, the coefficients of the variable *Progression* in Column III and VI represent the pure effect of a change in w. The result shows that with each additional year of progression enrollment increases by .9% and teacher/student ratio decreases by .8% caused by pure change in w.

Supply of High School Education

For high school education, a similar variable progression as in middle school case is not available since the nation did not promulgate a plan that makes high school education universal. In fact, by 2013 most provinces have upper-secondary school enrollment rate smaller than 85% and high school smaller than 60%. The high school entrance exam still exists and divide students with different scores into highly differentiated schools so the competition between high schools is still intense. Here I suggest to use the number of high schools per county as an indicator for weight on quantity (w).

There are two reasons that the number of high schools and weight on quantity have highly positive correlation. Firstly, the number of high schools may result in an increase w. If a county has more schools, it's more likely to form a market with schools having a variety of objective functions to accommodate students' demand. Therefore the proportion of quantity maximizers may increase, leading to a higher w. Secondly, a large amount of high schools in a county may be a result of higher w set by the government since more schools mean easier access to schools and relatively lower quality. Though the causality is hard to identify, the number of high schools can be an indicator of w.

Theres concern that the number of high schools per county could be affected by factors other than w. First, this variable is likely to be correlated with the level of education attainment or expenditure on education. For example, rich areas which set more premiums on education may invest to build more schools. Thus the effect of high school number reflects a difference in resources rather than that in the structure of high school system. To clarify this, I considered two variables: high school enrollment rate and average government appropriation per student in each province. Data shows that number of high schools per county is irrelevant with average enrollment level: the correlation between is 0.08. However, there is a moderate correlation of 0.4 between high schools per county and government

support on each student.

Secondly, more densely populated provinces are likely to have more high schools in each county. The correlation between school per county and population per county (high-school-age population per county) is as high as 0.60. It means that the variation in number of high schools per county is mainly determined by administrative division - in provinces which has larger population per county, an average county has more high schools to offer education to people all around the county.

Since school per county is highly correlated with average expenditure and population per county, I also use the two variables as a replacement for schools per county to check whether they have impact on enrollment or quality. As fixed-effect model cannot estimate coefficient on time-invariant variable, I add an interaction term between the three indicators for w and $LnGraduates$ to check how w affect elasticity. Its worth noting that the variable average expenditure could also serve as a weight on quality $(1-w)$ rather than on quantity(w) though it is positively correlated with schools per county. If average expenditure has opposite effect with w on enrollment and schooling quality, the choice of using the variable the number of high schools per county would be further proved to be appropriate.

I regress school enrollment (n) and teacher/student ratio (Q) on N, w, and g. The following table shows the result.

Table 3.9: Determinants of Enrollment and Quality in High School Education

	I	II	III	IV	V	VI	VII	VIII
Dependent variable:	LnEnrol	LnEnrol	LnEnrol	LnEnrol	Ln(Tch/Stu)	Ln(Tch/Stu)	Ln(Tch/Stu)	Ln(Tch/Stu)
Independent variable:								
LnGraduate	0.707***	0.563***	0.453***	0.735	-0.259***	-0.195**	-0.189*	-0.518***
	(0.0605)	(0.0720)	(0.0910)	(0.161)	(0.0426)	(0.0733)	(0.0946)	(0.119)
*LnGraduate**		0.0911*				-0.0407		
SchoolperCounty		(0.0525)				(0.0475)		
*LnGraduate**			0.366***				-0.101	
PopperCounty			(0.111)				(0.155)	
*LnGraduate**				-0.0434				0.396**
AverageExpenditure				(0.252)				(0.168)
Lng		-0.00473	0.0585	-0.0113		0.00517	-0.0112	0.0124
		(0.0806)	(0.0784)	(0.0812)		(0.0400)	(0.0384)	(0.0344)
Year dummy	yes	yes	yes	yes	yes	yes	yes	yes
Province dummy	yes	yes	yes	yes	yes	yes	yes	yes
N	390	390	390	390	390	390	390	390
R-sq	0.951	0.951	0.954	0.951	0.774	0.775	0.777	0.794

Note: * p<10% ** p<5% ***p<1%. The significance level on elasticity estimation is against the hypothesis that $\beta = 1$. Fixed effect model is estimated and robust standard errors clustered on province are reported in parentheses. Source: Chinese Educational Statistics Yearbook (1987-2013); China Educational Finance Statistical Yearbook (1987-2013)

The above table shows the regression result using different settings. The first Column shows that in general the enrollment elasticity is .71 in the period from 1998 to 2013. Column II added indicators for g and w and Column III and Column IV used different variables as indicators for w. Column V to VIII use schooling quality as dependent variable. It's clear that as the number of middle school graduate increases, enrollment increases with elasticity smaller than 1 and quality decreases. The model predicted that a rise in g will lead to an increase in n while its impact on quality is ambiguous. The empirical results show that as government resources increase, the priority is to increase enrollment rather than improving quality. As for w, the estimation is also consistent with model prediction. Column II and VI show that as N increases, a province with more high schools per enrollment unit would have more enrollment as well as lower quality. When population per county is used as an indicator for schools per county, the result is similar. Since there's no reason that a county with more population has a higher w, its impact on w should be through the indirect channel that a county with more population has more high schools and therefore higher w. Column IV and VIII shows that although financial means is positively correlated with number of high school per county, it's impact on enrollment and quality is totally different from schools per county. Therefore, schools per county affected enrollment decision through different weight on quality and quantity, not through educational resources available.

3.6 Hierarchies in High School: a Case Study in Shenyang

Nation-wide analysis using aggregate data is consistent with the idea that in high school quality is of more importance than in middle school. Further analysis on different types of high school (key-point school vs. ordinary school) can lend more support to the theory about quality-quantity tradeoffs.

Since I cannot find any national-representative data, I use data in Shenyang Prefecture from Yearbook of Examination and Admission in Shenyang (1991-2005). Shenyang is the capital prefecture in Liaoning Province, which contains 15 districts. Detailed information about high school entrance examination and admission is included in the yearbook. Basically therere two types of high schools: key point high schools and ordinary high schools. They follow the same curriculumn and students are selected from the same high school

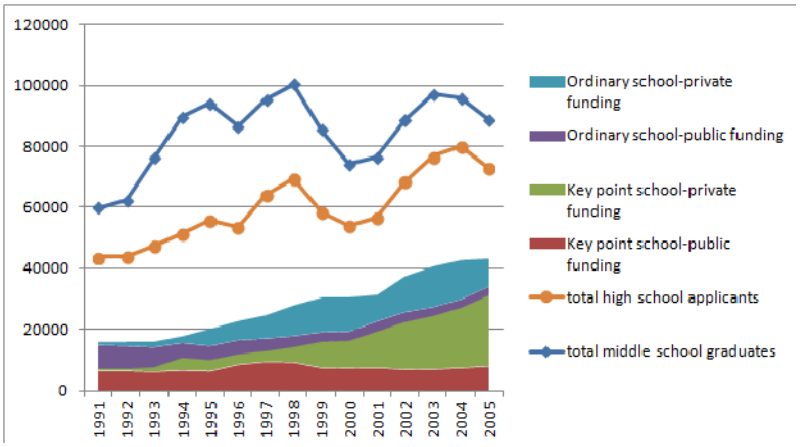
entrance exam and will attend the same college entrance exam, the only difference being key-point schools are deemed as better schools. In 2005 Shenyang has 52 key point high schools (28 provincial level key point high schools and 24 prefectural level key point high schools) and 53 ordinary high schools.

Just as all other areas in China after 1990s, most high schools in Shenyang have two different trajectories in admission: publicly funded admission and privately funded admission. Every year after the unified high school entrance exam, each school announces two different criteria: publicly funded criterion and privately funded criterion. Students whose exam score is greater than the publicly funded criterion pays a low tuition and students whose score is smaller than the publicly funded criterion but greater than the privately funded criterion can still enter the school by paying extra funding, which is called school selection fee. The"school selection fee" is usually ten times greater than the tuition. If a student narrowly met with a good schools privately funded criterion but not publicly funded criterion, he could either choose to pay extra money to enter the good school or become a publicly funded student in a lower ranked school with lower admission criterion. Also, its well understood that in the same school privately funded students may not receive the same level of education quality as publicly funded students: students with higher entrance exam scores are more likely to be enrolled in key point classes with better teachers and better peers. Generally speaking, the quality of education received is determined mostly by ones ability to score higher and partly by parents financial means.

The yearbook listed admission threshold by all key point schools and school selection fee by few key point schools in few years. For example, table 10 shows the criteria and charge by Shenyang No. 31 High school in 2002, which ranked 5th out of 105 high schools [32]. If a student scored more than 478 points (94% of total score), she could by admitted and publicly funded. As for privately funded students, this school has two different thresholds: if one scores less than 478 but more than 467 (92% of total score), she could be admitted by paying extra 5000 yuan as a school selection fee; if one scores less than 467 but more than 459 (90% of total score), she could be admitted by paying extra 15000 yuan. As a

[32] As official ranks change every year, I estimate instant ranking using admission threshold of public funded students

Figure 3.4: High school enrollment by type in Shenyang



reference, in 2002 the annual disposable income per capital of city residents in Liaoning province is 9730 and average expenditure on high school students is 4325 yuan[33]. A bit decrease in score means a big sum of money out of parents' pocket. To further illustrate the high demand on high quality education, in 2002 a student scored 466 can still enter a provincial key-point school which rank 18 out of all 105 high schools as a publicly funded student. Still, some parents would like to pay 12,000 yuan for a slot in the 5th ranked school.

Table 3.10: Criteria and charge in Shenyang No. 31 High school in 2002

Enrollment Type	Threshold	Payment (annual)
Public funded	478 (94%)	600 RMB
Private funded-1	467 (92%)	(600 + 5000/3) RMB
Private funded-2	459 (90%)	(600 + 5000) RMB

Source: Yearbook of Examination and Admission in Shenyang

Figure 4 shows the number of middle school graduates, number of high school applicants and number of enrollees by admission types from 1991 to 2005.

Theres huge fluctuation in the total number of middle school graduates, which is consistent with the national demographic change in birth cohort size. The fluctuation in number of high school applicants is concurrent with number of middle school graduates and the

[33]Though the school selection fee doesn't seem to cover an additional students' cost, it is a net income for the school.

proportion of application generally increased[34]. Also, the number of admissions more than doubled during this period due to the development of high school education. However, different types of admission showed different patterns. First, key point schools firmly refused to adjust its quota on publicly funded students. Theres very tiny change in the quota across the period. Second, privately funded students for both key point schools and ordinary schools increased a lot. In the years with relatively smaller number of applicants (1996, 1999-2001) the tempo of expansion slowed down. It could well be that schools use private-funding quotas to accommodate to varying demand.

In my basic model I did not consider endowments of student since a school is only able to attract students with a fixed range of endowments (see section 7 for further discussion)[35]. However, if students with different endowments pay different price, essentially the school regard students' endowment and expenditure as substitutes to produce a certain level of schooling quality. At the same time, the demand side would like to pay a higher price for a higher quality, forming a market equilibrium. This kind of market is essentially similar to America's college market, which forms strict hierarchies and charge different tuition according to ability and household income[36]. An increase in cohort size represents higher demand, therefore enrollment and school selection fee may increase at the same time. If the payment net of marginal cost is greater for private funded students than public funded students, the elasticity for private funded students will be higher than public funded students. Also, key-point schools, as quality maximizers, are more likely to restrain enrollment than ordinary schools although key-point schools can charge higher price for each slot[37].

---

[34]Middle school graduates who applies for vocational high schools do not count as applicants here as theyre enrolled through another system. The number is about 10,000 in early years and decreased as students seek for more education

[35]It's hard to say how endowment consideration affect schools' decision when cohort size changes. Intuitively, if a high school slightly increase admission slots from a large cohort, the average endowment may still rise, indicating higher promotion rate to college in the school's favor. However, if colleges do not increase enrollment accordingly, the relatively high endowment in a large cohort doesn't lead to a higher chance to be admitted by college and thus high schools do not have incentives to increase admission. Therefore I did not consider how students' endowment affect enrollment in my basic model

[36]see Epple, Romano and Sieg(2003); Winston(1999)

[37]In a similar analysis about American college, Bound and Turner (2006) found that top research colleges, including private and public, has lower enrollment elasticities than lowly ranked colleges

The yearbook provides data to examine the elasticity of different admissions. It has the number of applicants and enrollees for all 15 districts in Shenyang in 7 years from 1999 to 2005. With absence of detailed expenditure data, I still use the following regression similar to equation (5) to estimate the elasticity.

$$ln(Enrollment)_{idt} = \beta_i ln(Aapplicant)_{dt} + \delta_d + \rho_t + u_{idt}$$

where $Eenrollment_{idt}$ is the number of ith type of students enrolled in disctrict d in year t and $Applicant_{dt}$ is the number of total applicants to high schools in disctrict d in year t. District and time specific dummies are added to control for omitted variables. Four basic student types include public funded student in key-point school, private funded student in key-point school, public funded student in ordinary school and private funded student in ordinary school. The coefficients reported in the Table 11 are the elasticities for different types of enrollment.

Table 3.11: Enrollment elasticity by type

		Elasticity	**SD**
Key point school			
	public funding	0.189	(0.107)
	private funding	0.242	(0.102)
	total	0.257	(0.0711)
Ordinary school			
	public funding	0.721	(0.251)
	private funding	1.589	(0.476)
	total	1.088	(0.534)
Overall		0.604	(0.0656)

Note: Fixed effect model is estimated and robust standard errors clustered on district are reported in parentheses. Source: Shenyang Yearbook of Examination and Admission (1991-2005)

The elasticity of publicly funded enrollment in key point schools is as small as 0.189. This shows that to retain its stable quality on elite students, a key point school slightly changes its quota on publicly funded students when student number fluctuates. However, the elasticity on privately funded students is greater - partly due to more excessive demand of students from large cohort who are eager to pay extra money to enter better schools. Even

after counting extra payment students, the key point school elasticity is still small: a 10% increase in total high school applicants will only increase key point high school enrollment by 2.6%. The ordinary schools are more elastic. Its worth noting that elasticity for privately funded students in ordinary schools is as high as 1.59, greater than 1. It means 10% increase in high school applicants increase students in that trajectory by 15.9%. If extra students in a large cohort were refused by other three trajectories, the rest have to stand paying money to receive lower-quality education for they have no other choices. After all, the overall high school enrollment elasticity is .6, still smaller than 1.

As the elasticity of privately funded student in ordinary school is greater than 1, it challenges my basic assumption that demand is excessive than supply, at least not for this type of enrollment. It would seem that ordinary high schools want to admit more students and generate more revenue whilst families are reluctant to pay a high price for a relatively low return (as the probability of entering a good college is lower in ordinary schools). For those ordinary schools who're already at the bottom of hierarchies, we may assume they do not limit enrollment, ie. they are quantity maximizer, and their total admission is determined by parents' willingness to pay. Since key-point schools limit admission slots, the effective demand for ordinary schools multiplies as cohort size rises, resulting in an enrollment elasticity greater than 1.

It's worth noting that the policy on private funded students has been officially removed as Chinese government has sufficient financial means to fund high schools. In 2012, the new rule by minister of education capped the enrollment of private funded students in public high school to 20% of total enrollment and scheduled to remove private funded enrollment in three years[38]. All provinces gradually applied the new rule in following years. For Shenyang, the old regime was officially removed in 2014. Despite the egalitarian measures to neutralize the impact of household income on schooling opportunities, the disparity and hierarchy between key-point and ordinary high schools still exist and show no sign to be cancelled.

[38]source: Minister of Education (2012), Enforcement opinion on dealing with unregulated fees and charge by schools

3.7 Why hierarchies matter: a welfare analysis on demand side

As is shown in previous sections, in Chinese system there exist hierarchies in the same stage of education characterized by uneven resource distribution and limited enrollment. Intuitively, the hierarchies would be the source of education frenzy and high level of competition: students are not only competing for more education, but also for better education. They waste time on repeated practice to outdo their peers. In this section, I try to explain the behavioral mechanism of both schools and students using signaling theory and quantify the scale of loss in efficiency caused by competition.

In a certain year, N students graduate from lower-level school and want to be admitted by one of the m higher level schools in the administrative area. Suppose the endowment of students (e) follows a uniform distribution, $e \in (a, b)$ with $0 < a < b$. The utility of a student is determined by the utility of going to a specific school (the quality of the school) and her leisure time if going to school i, denoted as L_i ($L_i \in (0, 1)$).

$$U(L, i) = Q_i \times L$$

where $Q_1 > Q_2 > \cdots > Q_m > Q_0$, Q_i is the benefit of going to ith school and Q_0 is the benefit of attending no school. Suppose the score of the student in exam is S and it is a product of effort made and ones own endowment:

$$S = (1 - L) \times e$$

Therefore, the utility function becomes

$$U = Q_i \times (1 - \frac{S_i}{e})$$

As $Q_1 > Q_2 > \ > Q_m$ for all students, it hold true that $S_1 > S_2 > \ > S_m$. Denote e_i as the endowment with which a student feel indifferent between going to ith and $(i + 1)$th school, which means

$$Q_i \times (1 - \frac{S_i}{e_i}) = Q_{i+1} \times (1 - \frac{S_{i+1}}{e_i}) \tag{3.10}$$

Then for any student with $e > e_i$, she would strictly prefer ith school than $(i+1)$th school. The set of students endowments enrolled by each school is a continuum from high endowment to low endowment. Note that the school does not really care about the score; they care about the signal about endowment revealed in the score.

Suppose in an equilibrium, each school admits n_i students, it must be true that

$$\sum_{1}^{m} n_i <= N$$

and

$$\begin{cases} e_1 & = b - \frac{n_1}{N}(b-a) \\ e_i & = b - \frac{\sum_1^i n_i}{N}(b-a) \\ e_m & = b - \frac{\sum_1^m n_i}{N}(b-a),\ if \sum_1^m n_i < N; \\ & = a,\ if \sum_1^m n_i = N \end{cases}$$

given the score perfectly signals the endowments.

Suppose all students are fully informed about the enrollment quota of each school (n_i) and their own rankings of endowments among peers. S_i should be high enough such that students with $e < e_i$ would be worse off if she attains a score greater than S_i. Using equation (1) we have

$$S_i = \frac{Q_{i+1}}{Q_i} S_{i+1} + (1 - \frac{Q_{i+1}}{Q_i}) e_i$$

S_i is determined by the scores of schools with inferior notch.

So we start from the school with the lowest rank, we have the threshold

$$S_m = \begin{cases} (1 - \frac{Q_0}{Q_m}) e_m,\ if \sum_1^m n_i < N \\ 0,\ if \sum_1^m n_i = N \end{cases}$$

Therefore, we have

$$S_i = \frac{Q_m}{Q_i} S_m + \frac{\sum_{j=i}^{m-1} (Q_j - Q_{j+1}) e_j}{Q_i}$$

The above equations define an equilibrium that a students scores S_i and goes to ith ranked school to maximize her utility if her endowment e satisfy $e_i \leqslant e < e_{i+1}$, that each school

i is willing to enroll any student that has score greater than or equal to S_i, and that the highest ranked school believes its students' endowment ranked from 1 to n_1 and any other school i believes its students' endowment ranked from $\sum_{j=1}^{i-1} n_j + 1$ to $\sum_{j=1}^{i} n_j$. Therefore, for any given level of enrollment, the school is completely informed of the endowments of prospective students. The school system can also reach new equilibrium if it adjusts enrollment number and everyone is well informed. The mth school may reach a corner solution if $\sum_1^m n_i = N$, and an interior solution is feasible for all other schools.

Note that in equilibrium, a student with higher endowment than threshold endowment enjoys a certain level of welfare surplus: she would score just equal to threshold score rather than doing her utmost and therefore she would enjoy some free time and higher utility.

Therefore, people with utility e satisfying $e_i < e < e_{i-1}$ go to ith school and their total utility is

$$U_i = \int_{e_i}^{e_{i-1}} Q_i(1 - \frac{S_i}{e}) f(e) de = \frac{Q_i}{b-a}[\frac{n_i}{N}(b-a) - S_i(lne_{i-1} - lne_i)]$$

where $f(e) = \frac{1}{b-a}$ is the probability density function of e and define $e_0 \equiv b$. Thus total utility Θ for all students is the sum of those who went to higher school and those who didn't with utility equal to Q_0.

$$\begin{aligned}
\Theta =& \sum_{i=1}^{m} U_i + \int_a^{e_m} Q_0 f(e) de \\
=& \frac{1}{N}\sum_{i=1}^{m} Q_i n_i - \frac{1}{b-a}\sum_{i=1}^{m} Q_i S_i (lne_{i-1} - lne_i) + Q_0 \frac{e_m - a}{b-a} \\
=& \frac{1}{N}\sum_{i=1}^{m} Q_i n_i - \frac{1}{b-a}[\sum_{i=1}^{m-1} e_i(Q_i - Q_{i+1})(lnb - lne_i) + (Q_m - Q_0)e_m(lnb - lne_m)] \\
&+ Q_0 \frac{e_m - a}{b-a}
\end{aligned}$$

[39] Now consider a system without any hierarchy but the same enrollment in m schools, it

[39] $A \equiv \sum_{i=1}^{m} Q_i S_i (lne_{i-1} - lne_i) = Q_1 S_1 lnb - Q_m S_m lne_m - \sum_{i=1}^{m-1}(Q_i - Q_{i+1}) e_i lne_i$
As $Q_i S_i - Q_{i+1} S_{i+1} = e_i(Q_i - Q_{i+1})$ and $Q_m S_m = (Q_m - Q_0)e_m$, we have

$$Q_1 S_1 = (Q_m - Q_0)e_m + \sum_{i=1}^{m-1} e_i(Q_i - Q_{i+1})$$

means that all schools have the same quality $\bar{Q}$ and same threshold score $\bar{S}$. If we assume quality is just determined by resources offered by government, then essentially we could assume

$$\bar{Q} = \frac{\sum Q_i n_i}{\sum n_i}$$

The schooling quality in the new equal system is the weighted average of all schools' quality in the hierarchical system. Top ranked students now enjoy lower quality and bottom ranked students enjoy higher quality.

If we assume $\sum n_i = N$, i.e. no student will be left unenrolled, then in the equal system students would spend no time studying. There will be no welfare loss and utility is maximized.

But if $\sum n_i < N$, things become more complicated. Essentially students who are enrolled are exactly same students as in a hierarchical system. Among those enrolled, top-ranked students will spend less time studying as they face a lower threshold but bottom-ranked students are motivated to study more since $\bar{Q} > Q_m$ means $\bar{S} > S_m$ – they need score more to signal their ability against those not going to school as the difference in utility between going to school and not going to school increased. In this equal system, the total utility is:

$$\begin{aligned}\Theta' &= \int_{e_m}^{b} Q_i(1 - \frac{S_i}{e}) f(e) de + \int_{a}^{e_m} Q_0 f(e) de \\ &= \frac{1}{N}\bar{Q}\sum_{i=1}^{m} n_i - \frac{1}{b-a}(\bar{Q} - Q_0) e_m (lnb - lne_m) + Q_0 \frac{e_m - a}{b-a}\end{aligned}$$

The change in total utility caused by a hierarchical system is

$$\Theta - \Theta' = \frac{1}{b-a}[(\bar{Q} - Q_0) e_m (lnb - lne_m) - \sum_{i=1}^{m-1} e_i (Q_i - Q_{i+1})(lnb - lne_i) - (Q_m - Q_0) e_m (lnb - lne_m)]$$

To simplify the analysis I assume $\{Q_1, Q_2, ..., Q_m\}$ is an arithemetic sequence with common difference equal to D_Q, indicating the quality difference between schools with adjacent ranks. Also, each school admit n students, making $\{b, e_1, e_2, ..., e_m\}$ an arithemetic sequence as well

---

So $A = \sum_{i=1}^{m-1} e_i (Q_i - Q_{i+1})(lnb - lne_i) + (Q_m - Q_0) e_m (lnb - lne_m)$

and the common difference is $(b-a)\frac{n}{N}$. Therefore

$$\begin{aligned}\Theta-\Theta' =& \frac{1}{b-a}[(Q_m-Q_0)e_m(lnb-lne_m)+\frac{(m-1)D_Q}{2}e_m(lnb-lne_m)\\ &-\sum_{i=1}^{m-1}e_iD_Q(lnb-lne_i)-(Q_m-Q_0)e_m(lnb-lne_m)]\\ =&\frac{D_Q}{b-a}[\frac{(m-1)}{2}e_m(lnb-lne_m)-\sum_{i=1}^{m-1}e_i(lnb-lne_i)]\\ =&-\frac{D_Q}{b-a}\sum_{i=1}^{m-1}(\frac{blnb}{2}+\frac{e_mlne_m}{2}-e_ilne_i)\end{aligned}$$

[40]Since $\{b, e_1, e_2, ..., e_m\}$ is an arithemetic sequence and function $xlnx$ is convex, $\sum_{i=1}^{m-1}(\frac{blnb}{2}+\frac{e_mlne_m}{2}-e_ilne_i)>0$. Therefore

$$\Theta-\Theta'<0$$

This equation suggests that there is a loss in total utility due to hierarchical system. Also, note that D_Q, the difference in quality in the hierarchy, is determined by how government allocate resources and independent of the number of schools. Therefore D_Q and $\sum_{i=1}^{m-1}(\frac{blnb}{2}+\frac{e_mlne_m}{2}-e_ilne_i)$ are independent. Rises in m and D_Q both increase the loss in welfare. Therefore both number of hierarchies[41] and dispersion in quality will enlarge the loss in total welfare.

The result implies that the hierarchical and competitive secondary education system currently run in China is not favorable as it creates loss in efficiency. One may argue that the effort devoted to exam preparation is not a waste of time, only serving to signal ones ability, but rather a necessary cost to accumulate human capitals. However, since the secondary school system is exam-oriented in China, the curriculumn is monotonous and criticized as restricting students all-around development (e.g. Han and Yang, 2001). Lots of time was wasted on useless repetition or special techniques to improve scores, which can rarely be used in workplace or turn into monetary returns. Hartoga, Sun, and Ding (2010)

[40]$\sum_{i=1}^{m-1}e_ilnb = lnb[(m-1)e_m+\frac{m-1}{2}\frac{mn}{N(b-a)}]$

[41]A rise in number of school doesn't necessarily increase number of hierarchies; if m' school have same quality, i.e. same resources from government, it's equivalent to a system of $m-m'+1$ schools having different qualities.

found that self-rated quality of high school, while affecting quality of university attended, has no effect on earnings. Li et.al (2012) found that return to high school education is about 5%, much lower than return to upper-secondary vocational school education at about 20%. The low return in high school is staggering given the traditional belief that only incapable students who fail to enter high school go to vocational schools (Ling, 2015). It suggests that vocational school education provides training with market values while exam-oriented high school education has very limited effect on students human capital except selecting more capable students into college.

The hard working before college may also reduce students incentives to study in college, where real useful skills are gained. Lee (2007) pointed out that compared with American students who work harder in college to signal their ability, Japanese and South Korean students, who have already finished signaling process by entering big name colleges through working hard in high school, work less time in college. Though in China official data on college students studying time is absent, lack of training is also widely noticed for college students, leading to a public concern that the knowledge structure of college students does not satisfy the demand of the market (Yao, 2004). These results combined suggest that the hierarchical system has profounding negative impact on human capital accumulation and labor market performance.

3.8 Conclusion

In this paper I present evidence of cohort crowding in Chinas public school system. Children born in fat cohorts suffer not only fewer enrollment opportunities but also worse quality due to limited government budgets. Using census data, the estimation shows that in general a 10% increase in cohort size leads to 1% decrease in enrollment of middle school, 3-4% decrease in enrollment of high school and 4-5% decrease in enrollment of college. The negative impact is alleviated in recent decades, but richer information presented by institutional data shows more evidence on cohort-crowding, including smaller enrollment elasticities and negative effect on quality measures.

This paper also suggests a new angle on understanding the public provision of education

in China. With disparities in resource allocation and intense competition dominated by education bureaus, school authorities make tradeoffs between quantity and quality to serve their own end. Different schools responses to varying cohort size throw into sharp relief their different objectives. On being faced with high demand caused by big cohort size, high schools, compared with middle schools, are less likely to adjust admission slots and therefore less likely to sacrifice quality. Key point schools restrain their admission slots while ordinary schools are relatively flexible to adjust their slots. This paper also regards the implementation of compulsory education as a process that the whole nation puts more weight on quantity step by step and found evidence that middle school admissions become more responsive to changes in cohort size with time but failed to keep a constant quality.

Since no paper has studied the supply of education in China in such a quantitative way, this paper has important policy implications. My findings suggest that the strict hierarchy and disparity in resource allocation are the major reasons of inelastic supply and intense competition. Government should alter the exam-oriented hierarchical system in order to promote efficiency and equal enrollment opportunities for different cohorts. This paper also provides new materials for international comparisons. Contrary to China, in developed countries secondary education is universal and featured with low-competition. Despite its equity, the non-selective system is frequently criticized as leading to low education quality and reforms to add more selectivity are called for[42]. This paper, as a detailed analysis about a selective system, may provide insights for related policies in developed countries.

[42]Recently British government planned to expand its selective grammar schools aiming at "create an education system that will allow anyone in this country, no matter what their background or where they are from, to go as far as their *talents* will take them". This policy has aroused great controversies (see Scott (2016) for a summary)

Chapter 4

The College Application, Admission, and Matriculatio Decisions: Applications of Conditional Logit Model

4.1 Introduction

Students' transition from high school to college is a complex process that involves repeated interaction between students and colleges with both parties making series of decisions. In this process, there are at least three stages where the role of decision makers alter: first, students search for information from all sources and discuss with friends, parents and counselors about colleges they have interest in or potential to be admitted. They may choose several colleges strategically to which they send official applications. Second, colleges scrutinize applications submitted by students and decide to admit some of them based on their academic performance, college entrance exam, extracurricular activities, recommendation letter, and personal statement etc. Third, each student compares offers from different colleges and make the final matriculation decision. In this paper, I refer to the three stages as application - admission - matriculation decision respectively, and analyze the separate decision-making processes and the interplay between those processes.

In economics research, there is numerous literature that involves the college going process where students are selected into different colleges according to their characteristics and budget constraints (for example, Black & Smith, 2004; Dale & Krueger, 2002, 2014). Those papers seek unbiased estimation on how different variables affect one's probability to go to a college. However, research considering the college going process itself is relatively scarce. Due to lack of data, present literature either focuses on a special subset of students or colleges making either the matriculation or admission decision, or implicitly regard the whole college-going process with complex interaction between colleges and students as one single decision made by students, omitting constraints or limited choice set faced by them. The

first strand of literature focuses mostly on the choice of high-ability students or high-quality colleges using administrative data. Avery and Hoxby (2004) analyzed how high college aptitude student respond to scholarship and aid packages. Horstschraer(2011) analyzed how top-ability students respond to change in ranking of medical schools, assuming their choices are not constrained by admission threshold. Numberg, Schapiro and Zimmerman (2012) analyzed administrative record of a very selective college and studied factors affecting ones matriculation decision. Toutkoushian (2001) examines what factors influence a students decision to consider attending different types of institutions in the region, and whether student choice is affected by parental education and income. Weiler (1996) discussed how high-ability students choose between a certain selective college and other alternatives. In the later strand, the impact on students' enrollment decision of factors, including college quality, ranking, family background, tuition, and the distance between college and students hometown etc., has been explored (Hurwitz, 2011; Long, 2004; Monks & Ehrenberg, 1999; Tobias, 2002; Tobias, 2003). The problem of this strand is that it usually only focuses on the final matriculation decision without considering the alternatives faced by students. Therefore it becomes very difficult to distinguish the impact of a college's own characteristics from that of other colleges as information on other colleges are not available. Then omitted variable bias arises since the related variables cannot be considered exogenous with respect to the matriculation decision.

This paper fills the gap by analyzing the whole college going process comprised of application, admission and matriculation, taking use of the national representative dataset Education Longitudinal Study of 2002 (ELS:2002) which provides a large variety of variables recording students personal experience from high school to early work stage or further schooling. This paper contributes to existing literature by obtaining unbiased estimation on factors affecting college going process, taking advantage of complete information provided by ELS 2002 about the process from application to matriculation. This paper focuses on two separate decisions: admission decision and matriculation decision. We explicitly treat the two decisions as made by colleges and students separately. Using the conditional logistic model, the methodology in this paper allows for specific choice set faced by all decision makers and we have a large set of student-college matched variables to capture factors

considered by each party. Those factors include tuition, distance, and ability matching, etc.

Though we focus on colleges' admission decision and students' matriculation decision, the main contribution is made by adding information generated from application process. When considering the college's admission decision, several papers used the conditional logit model as it can eliminate college-specific fixed effect (for example, Hoxby and Avery, 2004; Klaauw, 2002; Long, 2004). However, it doesn't take care of unobserved student ability. This paper proxies unobserved ability using information revealed in application process by assuming that ability unobservable to econometricians is observable to colleges and students. Therefore students' applications to and admissions by colleges at different selectivity level can tell about the student's ability level. The trick here is that a college cannot observe an applicant's application and admission results at other colleges at all; therefore those variables will not affect the college admission decision directly. If the variables have impact on admission result after controlling for other observed characteristics, it should be through the channel that they are indicators of individuals unobserved abilities. Three sets of variables about a student's applications to colleges at different selectivity level are used: whether the student applied to colleges at one selectivity level or not, the number of colleges the student applied at one selectivity level, and the student's acceptance rate by colleges at the selectivity level. The result shows that for college at a given level, students who filed more applications to a higher level and less applications to a lower level are more likely to be admitted; higher acceptance rate by other colleges also predict higher probability of admission. The findings here also reveals students' behavior pattern in college application. They typically apply for some "reach" school, some "match" school and some "safety" schools according to perception about their own ability and chances to be admitted by different colleges.

This paper also casts light on college "matching". A burgeoning literature discusses about how low-income or other disadvantaged students select themselves into under-match colleges, i.e. less selective colleges that are below their abilities (Avery & Hoxby, 2013; Cortes & Lincove, 2016; Dillion & Smith, 2017; Smith, Pender, &Howell, 2013). However, the definition of "undermatch" is not very clear. The literature usually focuses on the gap between a student's rank in measurable academic performance and the rank of the college

she applied or enrolled and assumes such undermatch is an irrational choice. However, this paper suggests perhaps students apply to undermatched colleges based on their rational expectation about their chances to be admitted. For example, Avery and Hoxby (2013) pointed out that most under-matched high ability students are from districts too small to have selective public high schools and therefore the lack of application is due to lack of information. But our result may suggest another story: due to the lack of exposure to high ability peers or teachers, those students may have disadvantage on other unmeasurable factors like recommendation letter, personal statement etc. and thereby choose not to apply to selective colleges with concerns on their chances to be admitted. Therefore, interventions like more counseling or preparation, rather than simply "informing" or "reaching" those students, might be of greater help.

The remainder of this paper proceeds as follows. Section 2 presents introduction about the dataset we use and some descriptive statistics. Section 3 elaborate on the general college going process and analyzed the application patterns shown in data. Section 4 discusses how colleges make decisions, focusing on addressing unobserved ability problem using information revealed from application process. Section 5 discusses factors affecting students' matriculation decision using conditional and nested logit model. Section 6 is conclusion.

4.2 Data and Descriptive statistic

We use a national representative dataset Education Longitudinal Study of 2002 (ELS2002) conducted by National Center for Education Statistics (NCES). The survey sampled 750 schools nation-wide and interviewed over 15000 randomly selected 10th graders in 2002. Those respondents were then re-interviewed in 2004, 2006, and 2012. Like other educational longitudinal survey commissioned by NCES, the ELS2002 data provides a large variety of variables recording students personal experience from high school to early work stage or further schooling. Also, this dataset has its own advantage of containing a full record of each students college application, admission and matriculation results, specifically contained in 2006 wave as most college-going students are supposed to have entered college in 2006. In the second follow-up, students were asked to report all colleges they have applied to when they graduated from high school, as well as the application result and their final

matriculation decision. he data file contains 16,197 respondents, among which 11373 filed at least one application to colleges. In total, they filed 30402 applications to college.

The application data is combined with Integrated Postsecondary Education Data System (IPEDS) to merge key variables about colleges. IPEDS provides various information about each college every year, including but not limited to enrollment size, student to faculty ratio, 25/75 percentile of SAT/ACT score, acceptance ratio, tuition and financial aids etc. Since students have general perception about the "quality" of a college, by which they make decision about whether to apply/enroll or not, and colleges of different types may have different admission standards, it's important to categorize colleges in conformity with their reputation, attractiveness to students and academic standards. I appeal to Barrons Admission Competitiveness Index Data as the basis for classification. The 2004 series were used as it was the latest year before students make decisions in around 2006. Following Barrons college admissions competitiveness ratings, all four-year colleges in the data are categorized into five groups: 1 Most Selective, 2 Highly Selective, 3 Very Selective, 4 Selective, 5 Not Selective [1].

Table 1 shows descriptive statistics by students application situations. Column I uses the full sample of respondents in the dataset, Column II contains the subset of respondents who filed at least one application to colleges, Column III contains the subset of respondents who eventually attended four-year colleges, Column IV contains the subset of respondents who only applied to four-year colleges[2]. The statistics include students individual characteristics, family background, characteristics of high school, and their applications to different kinds of colleges.

Table 1 shows that compared to non-applicants, a larger proportion of college applicants are female, Asian or White, and more likely to come from family with higher socioeconomic status. Compared to a larger sample of having applied to at least one college, students who eventually went to four-year colleges or only applied to four-year college significantly have better family background, are more likely to come from private high schools, have better

[1]Barrons Index contains 7 categories. For simplicity, I combine the last three categories Less Selective, Nonselective and Special into Not Selective

[2]Students who applied to for-profit colleges were also dropped (52 students were dropped), as were the for-profit colleges themselves

Table 4.1: Descriptive Statistics

	I	II	III	IV
N	16197	11373	6375	6669
Female	50.1%	54.0%	54.7%	54.1%
Race				
Asian	9.0%	9.7%	11.0%	11.3%
Black	12.5%	12.0%	10.0%	10.5%
Hispanic	13.7%	11.9%	7.9%	8.4%
White	53.6%	56.4%	62.1%	60.5%
Other	5.3%	5.1%	4.7%	4.8%
Family Income				
<$25,000	21.0%	17.3%	11.5%	12.0%
>$25,000, <$50,000	30.4%	28.8%	23.9%	24.4%
>$50,000, <$100,000	33.9%	36.3%	40.0%	39.5%
>$100,000	14.8%	17.6%	24.6%	24.2%
Mother's Education				
Below High	12.0%	9.0%	5.2%	5.2%
High School	25.4%	22.9%	18.7%	19.0%
Some College	31.3%	32.9%	31.2%	31.7%
Bachelor	17.4%	20.2%	26.1%	25.7%
Master or Above	8.5%	10.3%	14.6%	14.1%
High School Characteristics				
Public	75.2%	71.3%	62.9%	63.8%
More than 50% graduates in 2003 attending 4 yr colleges				
	95.8%	87.2%	89.0%	89.4%
Percentage of students taking AP course				
	17.4	17.9	19.4	19.5
	(13.8)	(14.2)	(14.9)	(14.8)
Test				
Taken SAT	35.7%	45.0%	61.1%	60.3%
Taken ACT	35.7%	44.1%	53.8%	52.1%
Ave SAT Equiv.	1005	1019	1079	1076
	(208)	(206)	(188)	(190)
# of Applications to				
Most selective	0.199	0.283	0.482	0.470
	(0.843)	(0.994)	(1.28)	(1.258)
Highly selective	0.209	0.297	0.480	0.475
	(0.621)	(0.723)	(0.891)	(0.879)
Very selective	0.365	0.520	0.805	0.795
	(0.776)	(0.882)	(1.024)	(1.013)
Selective	0.506	0.721	1.000	0.987
	(0.892)	(0.989)	(1.073)	(1.067)
Not selective	0.246	0.351	0.429	0.428
	(0.576)	(0.66)	(0.725)	(0.717)
Total 4 yr colleges	1.525	2.171	3.195	3.156
	(2.079)	(2.18)	(2.192)	(2.166)
Total 2 yr colleges	0.352	0.502	0.084	
	(0.613)	(0.679)	(0.309)	

SAT/ACT performance, and filed more applications to four year colleges.

To focus on decision making process about four-year college, I limit my analysis to individuals who only applied to four-year colleges. Therefore the current analysis is representative for students who have relatively firm intention to acquire at least four-year postsecondary education.

The rich data set allows us to use a wide variety of variables that affects college enrollment decision to account for heterogeneity. We group those variables into the following three categories: student characteristics, college characteristics, and student-college matched variables. The sets of variables used in different decisions are different.

Individual characteristics include gender, race, family income, mothers education, and extracurricular activities. College characteristics include tuition, enrollment size, whether the college is public or not, and indicators in Barron's competitiveness level, generated from IPEDS 2006 data and Barrons Competitiveness index 2004. Since IPEDS provide tuition charges on in-state and out-of-state students, the variable TUITION takes the residency of students into consideration. Matched characteristics between the student and the college focuses on the admission package offered to students. Although the real tuition payment net of merit aids and financial aids are not available, the ELS2002 dataset included dummy variables on types of financial aid one receives, including scholarship, student loans, and student working opportunities. Those dummies combines with tuition can help us understand how financial factors affect students choice. In matched variables also included are other two important factors: distance and SAT score. Its plausible that students are more likely to choose a college that is within commuting distance since it allows the student to live at home to reduce her financial burdens. In each student - college pair, a dummy of living within 50 miles from the college is added to control for this effect. Also, if the college one applied to is too far away from the students hometown, the probability of attending the college will be lower but the marginal effect of an extra mile may be negligible. Therefore another dummy of living more than 200 miles away from the college is added. As for academic performance, human capital theory suggests that it would be most efficient if a student select herself into a college that matches her academic ability. Although students may prefer colleges with stronger peers and there is evidence that students' performance

Table 4.2: Independent Variables

Student Characteristics	
Female	A dummy equals to 1 if gender=female
Race	Categorical variable: 1 Asian; 2 Black; 3 Hispanic; 4 White; 5 Other
Family Income	Categorical variable: 1 <$25,000; 2 >$25,000, <$50,000; 3 >$50,000, <$100,000; 4: >$100,000
Mother's Educ	Categorical variable: 1 High school or below; 2 Some college; 3 Bachelor or above.
SAT math	percentile of SAT math score
SAT verbal	percentile of SAT verbal score
Extracurricular	A series of binary variables on whether one has attended extracurricular sports, music, academic, volunteer and other activities in high school or not.
College Characteristics	
Tuition	Tuition in 2006 ($1000). In-state tuition applies if individuals residency in first follow-up is in the same state with the college.
Enrollment	Number of enrollment (1000) in 2006.
Public	A dummy variable equals to 1 if the college is public.
Barron's Selectivity	Categorical variable using Barrons college competitiveness index: 1 Most selective; 2 Highly selective; 3 Very selective; 4 Selective; 5 Not selective
Student-college match	
Grant	A dummy variable equals to 1 if one received scholarship or tuition rebate from the college
Loan	A dummy variable equals to 1 if one received student loan from the college
Student work	A dummy variable equals to 1 if one received student working opportunities from the college
Distance_community	A dummy variable equals to 1 if one lives within 50 miles from the college (commuting distance)
Distance_faraway	A dummy variable equals to 1 if one lives 200 miles away from the college
Sat Difference	Equals the absolute value of colleges median sat percentile minus the students own sat percentile
Sat Difference_plus	Equals 1 if the value of colleges median sat percentile minus the students own sat percentile is positive; 0 if otherwise.

improve with better roommates or small peer groups, studies also show that in a more competitive program students are less likely to obtain higher GPA or graduate (See Luppino and Sanders (2015) for a summary) [3]. In order to measure the matching quality, I calculate the absolute value of the difference between a colleges median SAT percentile and the applicants' own SAT percentile[4][5]. Since a positive value and a negative value may have different marginal effect on matriculation chances, I added a dummy on whether the difference is positive.

4.3 College going process: an overview

To model the college enrollment process, Long (2004) suggests the college decision consists of two choices: the individual determines his best college and concurrently decide whether to attend college or not at all. Due to the lack of data, the first step is simply a mapping from potential colleges one may have interest in to the college he eventually enrolled. However, its understood that college enrollment is a multi-stage process where different parties (college and student) make decisions based on alternatives they have, rather than choosing from an infinite set of alternatives. Therefore, we suggest to follow the natural process of college enrollment and model different stages separately following Manski and Wise (1984).

In general, the college enrollment outcomes are the result of the following three sequential decisions:

1. Students choose the colleges theyre interested in and apply to those colleges.
2. Colleges make admission and scholarship decisions.
3. Students make final matriculation decision among all colleges that have accepted her[6].

The question about the first decision, i.e. the application decision is to understand how

[3]Students are also supposed to make choice on their majors which may significantly affect their choice on colleges. However, such further analysis is not allowed since data did not provide information on majors associated with students' application

[4]Data Source: *SAT Percentile Ranks for Males, Females, and Total Group*, College Board 2006

[5]Since the IPEDS data only provides 75th and 25th percentile SAT score of students enrolled, I used the average of them to proxy the median SAT score

[6]In this paper we deliberately overlook the college/work choice since our focus is on how students choose college with different characteristics. To simplify the analysis, we limit the analysis to the behavior of people who only applied to four-year non-profit colleges, assuming theyre to a certain degree more determined to go to any college rather than working as a high school graduate.

students choose one or a few colleges among thousands of colleges in the U.S. This at the first glance may look like a typical choice problem but the practical question is how to define the choice set. Avery& Hoxby (2013) suggest to regard ALL colleges in the U.S as the choice set for each student. Under this setting, their usage of random utility model implies that all colleges to which student actually applied must incur greater expected utilities than any other college. There are two issues with this very strong assumption. First, the assumption about utility between chosen and unchosen colleges does not comply with the common understanding about student behavior. Usually students have an evaluation about their competitiveness and choose colleges according to their competitiveness level, with some consideration on other school and individual characteristics. In most cases students choose not to apply to a very selective college not because her expected utility is lower but because her chances to be admitted are remote, which can hardly be captured by variables used in estimation. Also, as the full set of colleges contains too many colleges usually a student does not have information about all choices available. They simply search information about colleges they have heard of and make choices among them. Second, in the application stage students usually choose several colleges strategically in a stratified manner. Counselors or reports guiding students college application always suggest them to follow some tactics: to guarantee a good match, a student should apply to several match colleges on par with her ability level (usually measured by average SAT, or by personal experience of previous students); to try her chance at better colleges, the student should also apply to some reach colleges which she has low probability to be admitted; to secure a position to pursue a bachelors degree, one or two safety college that the student is almost guaranteed admission is recommended. Therefore, different alternatives that a student eventually choose serve for different ends, i.e., the choice of one college is not independent of other choices made by the student. Therefore, its important to build linkage between different colleges chosen by a student in order to take full information into consideration.

With all the shortcomings of modeling each binary decision of applying to a certain college separately, however, its hard to find a way to model the joint decisions to apply to several colleges. Therefore we suggest to circumvent this problem temporarily and just use application information as an instrument for self-evaluation when thinking about admission

and matriculation decisions. The following part in this section just takes a preliminary look at the information shown in application process.

Table 3 lays out a brief comparison between national enrollment data in year 2006 and the final enrollment in ELS2002 data. For all colleges that I have complete enrollment information, the left panel shows the distribution of total enrollment numbers in 2006. The right panel shows students enrolled by different colleges in ELS2002 data. The two sets of data are pretty consistent in terms of share of enrollment by college selectivity level.

Table 4.3: Number of enrollment to different type of colleges

	College enrollments			Student sample.		
Barron's index	total	perc.	cum.	total	perc.	cum.
Most selective	90890	6.5%	6.5%	478	7.6%	7.6%
Highly selective	152920	11.0%	17.5%	720	11.4%	19.0%
Very selective	342609	24.6%	42.1%	1620	25.7%	44.7%
selective	553371	39.7%	81.9%	2512	39.8%	84.4%
Not selective	252532	18.1%	100.0%	982	15.6%	100.0%

To show the general application patterns, Table 4 exhibits applications to all kinds of four year colleges by the selectivity level of the college that a student eventually matriculated to in our sample consisting of students who only applied to four year colleges.

In table 4, rows represent the type of colleges a student eventually matriculate to and columns represent corresponding students' application behaviors. For example, in the sample 457 students eventually matriculated to most selective colleges. All of them applied to most selective colleges; on average they applied 3.36 most selective colleges; 269 of them applied to highly selective colleges and on average 1.68 applications are filed. The table shows that a large proportion of students filed applications to a variety of colleges with different selectivity level. Students matriculated to a given level are most likely to file applications to the level itself and adjacent selectivity levels. There are more applications filed especially for those who intends to go to most selective colleges.

Table 4.4: Student application by type of colleges matriculated to†

Matriculated to	# Students	Also applied to									
		Most selective		Highly selective		Very selective		selective		not selective	
Most selective	457	457	(3.36)	269	(1.68)	223	(1.45)	105	(1.36)	27	(1.3)
Highly selective	664	301	(2.13)	664	(1.82)	379	(1.61)	244	(1.46)	58	(1.07)
Very selective	1464	297	(1.64)	452	(1.48)	1464	(1.75)	603	(1.52)	186	(1.16)
Selective	2049	155	(1.42)	347	(1.3)	721	(1.39)	2049	(1.78)	409	(1.28)
Not selective	1078	41	(1.32)	102	(1.14)	224	(1.29)	437	(1.44)	1078	(1.38)
Not enrolled by any	938	115	(1.64)	201	(1.3)	367	(1.38)	567	(1.56)	418	(1.26)

† Numbers in parenthesis represent average number of applications filed.

Figure 4.1: Probability of applying colleges in each selectivity level

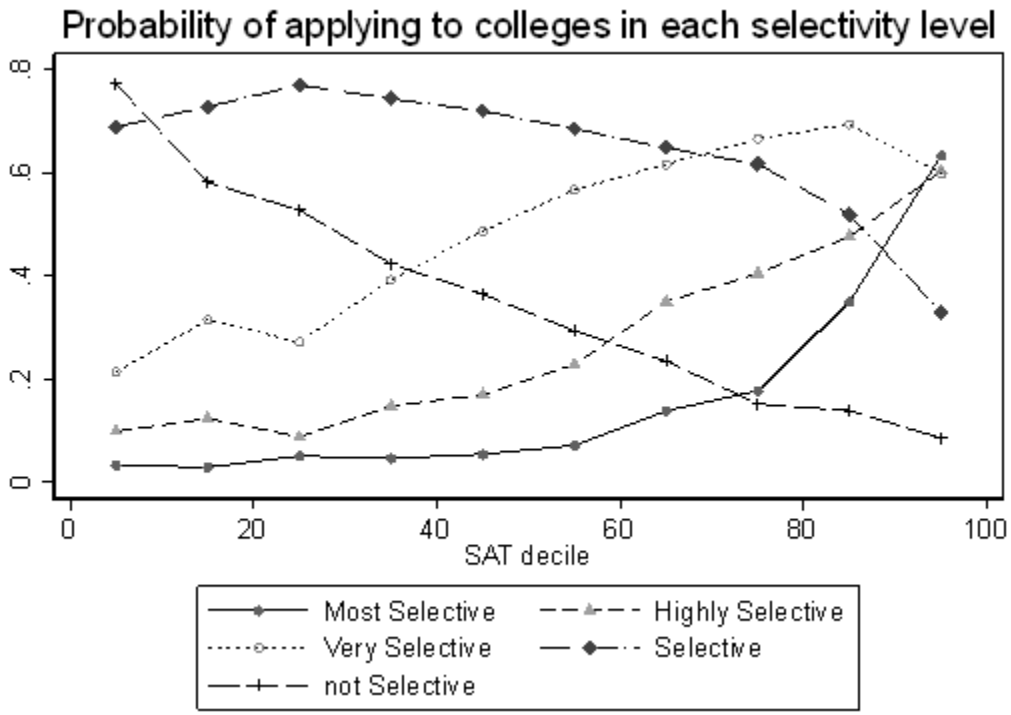


Figure 4.2: Number of colleges applied to in each selectivity level

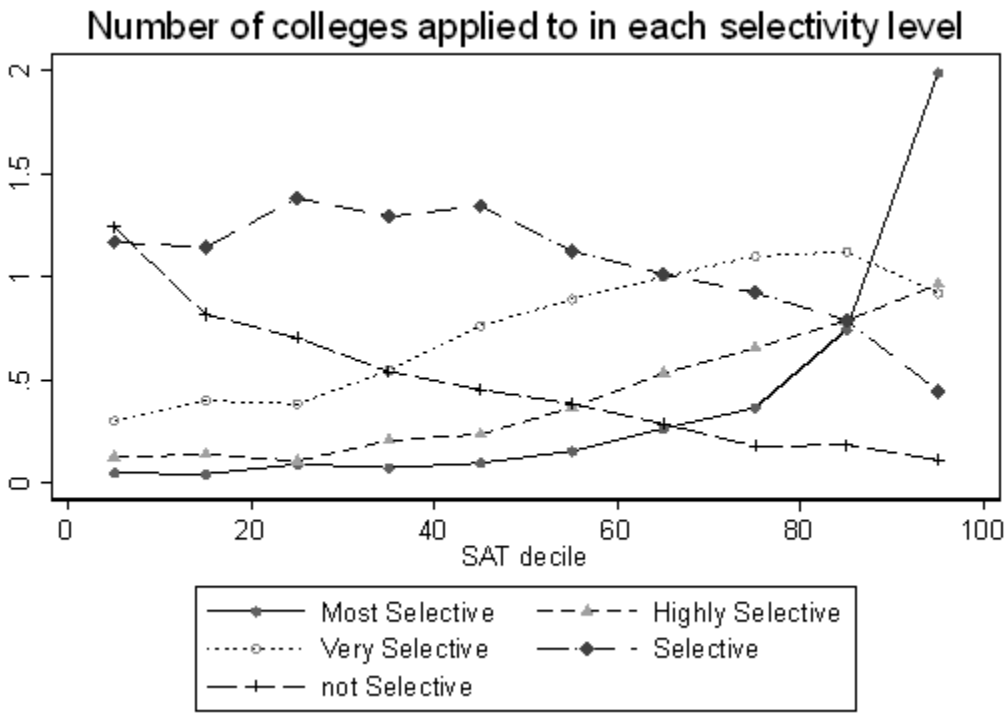
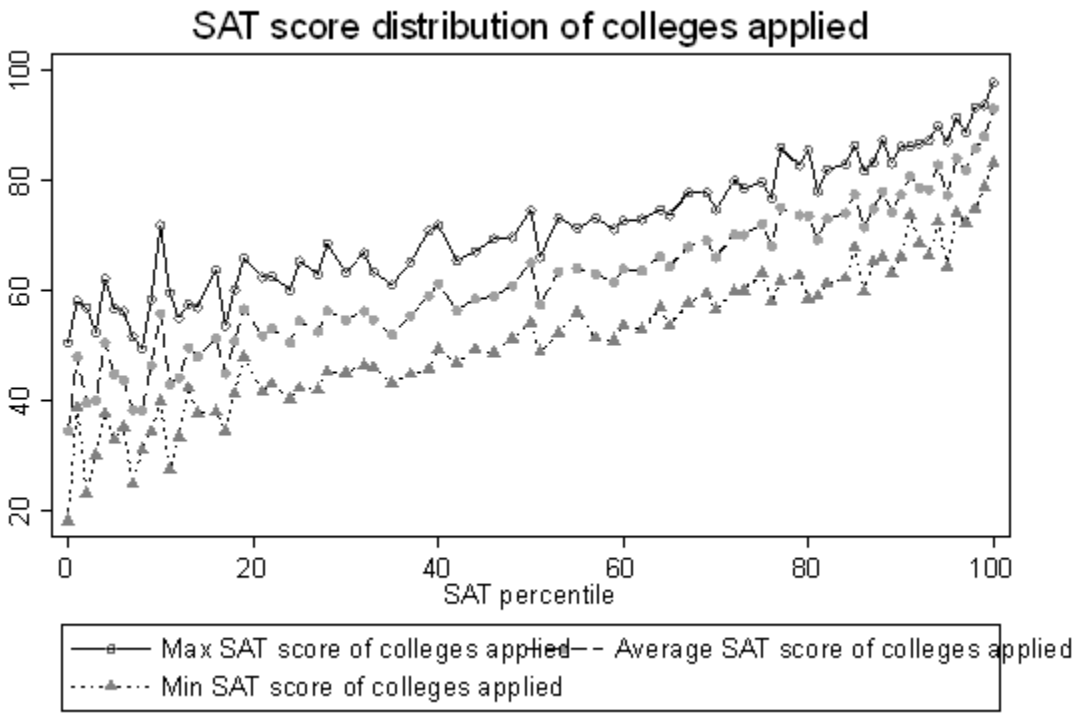


Figure 4.3: SAT score distribution of colleges applied to



To further see the linkage between college application behavior and one's academic ability, I use SAT score as a proxy for academic ability. I generate a dummy variable by categorizing students into ten SAT score deciles and run regressions on whether a student applied to colleges in a certain selectivity level using logit model, controlling for other variables including gender, race, family income, mother's education, GPA, and extracurricular activities. (Note this is a preliminary estimation without considering any endogeneity, the relationship shown is pure correlation rather than causality.) Using the estimated coefficient, figure 1 shows the probability that a student with a given SAT decile apply to different kind of colleges given an average value of other variables for students in the corresponding decile. For colleges in the first two selectivity levels - most selective and highly selective, the probability of applying increases as the student has higher SAT score. For the third selectivity level, the probability of applying also increases with SAT score except for the highest decile, since students with highest SAT scores has high expectations on being admitted by better colleges. The probability of applying to the fourth selectivity level is inverse U-shaped, with both low and high ends of SAT distribution having lower probability. For colleges in the lowest selectivity category, students who have rankings on par with the college are most likely to apply and the probability decreases with SAT score. For the top 10% students, their probability of applying to most selective, highly selective and very selective colleges are very similar. Though they're supposed to match into most selective schools, they still apply to highly selective and very selective colleges with high probability, reflecting their strategic behaviors when arranging application.

Figure 2 shows the number of colleges applied to in each selectivity level by SAT deciles (Depedent variable equals 0 if applied none in that category). In general, top 10% students filed the most applications. Clearly they target at most selective colleges by filing in average two applications to those colleges. Students at 80-90 percentile SAT score filed the second most applications to selective colleges but each filed only about 0.68 applications. Also, although their SAT percentiles "match" with "Highly Selective" colleges, they file most applications to a lower level, "Very Selective" colleges.

Figure 3 shows the relationship between a student's own SAT percentile and SAT percentile of the colleges she applied to. On average, the best college one applied has SAT

score 10 percentiles higher than average colleges a student apply to; similarly, students also apply to colleges with SAT score 10 percentiles lower than the average colleges as their "safe" colleges.

4.4 College admission

This section discusses the second stage of college going process: how colleges make their decisions to admit students among all applicants. The trick here is that though ELS2002 is designed to study the behavior of individual students, we could use the data to study colleges' admission decision. If we assume the student body is randomly sampled and they report all colleges they applied to, the data also reveals national representative colleges' admission behavior. In total, 1281 four-year colleges appeared in the data; 1122 of them received more than one application and 697 of them have both acceptance and rejection records.

4.4.1 Methodology

Suppose college c will admit individual i only when the observed and unobserved variables combined exceed the admission threshold of the college

$$Ad_{ic} = 1\{\beta X_{1i} + u_i + \epsilon_i > \theta_c\}$$

i.e.

$$Ad_{ic} = 1\{\beta X_{1i} + u_i + \epsilon_i - \theta_c > 0\}$$

where X_{1i} is a vector of observed characteristics of student i, u_i is unobserved characteristics of student i, and θ_c is the threshold of college c. We also assume ϵ_i follows i.i.d extreme distribution to apply logit model.

Obviously, the characteristics of applicants is correlated with a colleges threshold. i.e. $Corr(X_i, \theta_c) \neq 0$. With a panel data setting in which each college is matched with several individuals, a natural way to deal with unobserved threshold is to regard θ_c as college fixed effect using a series of dummy variables. However, unlike a continuous model, the inclusion

of a series of dummies will lead to incidental parameter problem and produce inconsistent estimates when the number of applications to each college is not big enough. Therefore, we use conditional logit model, which is based on a transformation of the problem that eliminates the fixed effect, to produce a consistent result.

The second source of bias in the estimation is unobserved individual characteristic u_i.[7]. This error term should capture student's characteristics like recommendation letters and personal statement. Note that although u_i is not observable to econometricians, it should be observable to the colleges since colleges make their decisions based on those factors. They may include subjective evaluations on students recommendation letters, personal statement etc. Therefore the admission decision made by different colleges may provide inference on the students ability. Meanwhile, though u_i are factors affecting the colleges decision and the standards are set by the college, the student should also have an estimation on it since she prepared all application materials. If we assume the student is rational, colleges she applied, perhaps including reach, match, and safety, should reveal her self-estimation on her own advantage.

The seminal research that incorporate college application and admission results to solve potential bias caused by endogeneity is Dale and Krueger (2002). In this paper, they used data from College and Beyond (C&B) survey where respondents were asked to list four colleges which they have seriously considered and the results of application. They used two methods to solve potential endogeneity: 1. The matched-applicant method, which assumes students who applied and accepted by similar colleges should have similar unobserved ability. In the estimation, fixed effect is added for similar pairs. 2. The self-revelation model which assumes that students know more about her own ability so colleges she has applied to reflect his true ability. In the estimation, the ability error term is proxied by average SAT score of all colleges she has applied for. In our binary outcome estimation for admission, the first method will produce inconsistent result as incidental parameter problem arises; the second

[7]Essentially there are two sources of endogeneity in this model: unobserved individual error term and unobserved college error term. The application of conditional logit model allows us to eliminate either of the two error terms as many individuals applied to several colleges and many colleges received several applications in the dataset. Hurwitz (2012) chose to eliminate individual error terms using conditional logit model and add college fixed effect as the number of applications to each college is big enough ($T \approx 10,000$) to produce unbiased estimation. In our setting, however, neither dimension produces large enough T.

method simply assumes that those who applied to a same set of colleges have the same abilities, omitting the probability that they may have different admission outcomes.

In this paper we suggest to use a set of variables describing ones application choice and admission result as proxies for a students abilities that are observable to the student and colleges she applied for but unobservable to econometricians.

$$u_i = \sum_{m=1}^{5} \{\gamma_{m1} DummyApplied_{im} + \gamma_{m2} NumberApplied_{im} + \gamma_{m3} AcceptanceRate_{im}\}$$

where $m = 1, 2, 3, 4, 5$ represents five categories in Barrons college selectivity level. $DummyApplied_{im}$ is a dummy whether one applied to any college in Barrons ith selectivity level or not, $NumberApplied_{im}$ is the number of colleges in Barrons ith selectivity level that one has applied to and $AcceptanceRate_{im}$ is the acceptance rate. The trick here is that the variables cannot be observed by any college at all; therefore they will not affect colleges admission decision directly. If the variables have impact on admission result after controlling for other observed characteristics, it should be through the channel that they are indicators of individuals unobserved abilities.

There is only one way, as far as I know, that colleges can observe applicants' applications to other colleges. It's a regular practice for colleges that in processing financial aid application, students are asked to file Free Application for Federal Student Aid (FAFSA), a unified national system to determine students' eligibility for different kinds of financial aid and facilitate financial aid provision. During the application, students are given the option to indicate up to 6 colleges (in 2002) where they wish to have the FAFSA data sent. Eventually a summary report, Student Information Report (ISIR), for each student will be produced and sent to colleges and the admission offices will be able to see other colleges the student sent ISIR to, as well as the order listed by the student. Though there's no research about how the information revealed by FAFSA affected college decisions, in 2014, a report by an admission consultant says that such FAFSA data might be used strategically by admission office (Graber, 2014).

Due to rising concerns by the public, FAFSA has determined to stop revealing such information to colleges starting in 2016-2017. Although the related report did not provide

strong evidence on how colleges utilize FAFSA information[8], in the empirical analysis I checked how FAFSA variables affect admission result.

One major problem for using conditional logit model is that the independence from irrelevant alternative (IIA) assumption may not hold. This assumption requires that the addition of one alternative does not affect the decision makers choice between existing alternatives, which in our case means the existence of one applicant does not affect the college's admission decision on others. Given the fact that many universities call forth diversity in student body and adopt affirmative action, one may assume the admission decision will be affected by the components of applicants, for example, with one additional Asian applicants, the probability of other Asians being admitted will decrease but that of students of other ethnicities will not change. This should not be a concern since in practice colleges does not explicitly set quotas for students with different socioeconomic background; instead the preferential treatment to disadvantaged takes the form that the college adds the test score of them by a fixed number of points (Pastine & Pastine, 2012). Therefore, the preferential treatment could be captured by adding control variables without harming the IIA assumption. In empirical part, I will also test whether IIA holds by dropping white applicants.

4.4.2 Empirical Results

We assume colleges in different categories may have different admission standards. Therefore we run the estimation separately for colleges with different selectivity. The following tables show conditional logit results. Table 5 uses the traditional setting and table 6 added our proxies for unobserved individual characteristics[9].

---

[8]Using administrative data from 153 colleges, the research referred by this report found that the ranking revealed by students have strong correlation with enrollment rate. The students enrollment rate is 64% for the first-ranked college and 22% for the second-ranked college and become lower for lowly-ranked college. However, the higher enrollment rate means students are more likely to attend higher-ranked college rather than the college is more likely to admit students who ranked them higher. Also, there are reason to doubt this report suffers under-report bias since their administrative data must include acceptance rate associated with different preferential orders but they did not report it.

[9]In order to make the variable acceptance rate exogenous to the dependent variable, in each regression the individuals who applied to only one college in that category are dropped, making the sample in the regression different from the full sample. To make the result comparable, in the traditional setting I also dropped those who applied to only one college in each category

Table 4.5: Conditional Logit on College Decision

	I Most Selective	II Highly Selective	III Very Selective	IV Selective
Independent variable:	1=Accepted; 0=unaccepted			
Female	-0.316**	-0.360**	0.114	0.338**
	(0.123)	(0.181)	(0.151)	(0.134)
Race				
Asian	0.00364	0.560**	0.329	-0.298
	(0.142)	(0.264)	(0.234)	(0.223)
Black	2.473+	0.661*	0.591*	-0.213
	(0.323)	(0.351)	(0.344)	(0.194)
Hispanic	1.191+	0.878+	0.266	0.105
	(0.268)	(0.314)	(0.304)	(0.234)
Other	0.771+	0.276	-0.107	-0.176
	(0.297)	(0.387)	(0.298)	(0.286)
Family income				
2	-0.445*	-0.382	-0.0324	0.0757
	(0.267)	(0.364)	(0.312)	(0.200)
3	-0.495*	-0.413	0.0256	-0.138
	(0.264)	(0.360)	(0.298)	(0.203)
4	-0.342	-0.0807	-0.100	0.0359
	(0.273)	(0.369)	(0.310)	(0.237)
Mother's education				
2	0.0224	0.385	-0.210	0.228
	(0.219)	(0.274)	(0.213)	(0.153)
3	0.235	0.0340	-0.332	0.346**
	(0.192)	(0.250)	(0.212)	(0.166)
SAT Math	0.00469+	0.00444+	0.00439+	0.00615+
	(0.000953)	(0.00128)	(0.00106)	(0.000920)
SAT Verbal	0.00417+	0.00511+	0.00323+	0.00285+
	(0.000805)	(0.00108)	(0.000951)	(0.000853)
GPA	1.512+	1.576+	1.991+	1.128+
	(0.186)	(0.198)	(0.180)	(0.136)
Taking AP course	0.364**	0.268	0.0631	0.314
	(0.172)	(0.190)	(0.178)	(0.197)
Extra curricular				
Sports	0.0166	0.301	0.214	-0.0742
	(0.132)	(0.189)	(0.171)	(0.148)
Music	-0.0261	-0.0315	0.125	-0.153
	(0.118)	(0.173)	(0.160)	(0.139)
Academic	0.125	-0.0524	0.00347	-0.219*
	(0.191)	(0.195)	(0.154)	(0.132)
Other	-0.0133	0.0944	0.243*	0.209
	(0.133)	(0.166)	(0.145)	(0.128)
Hours	0.0295	0.0882*	0.0769	0.0470
	(0.0381)	(0.0530)	(0.0479)	(0.0400)
Volunteer	0.343	0.614**	0.222	0.122
	(0.233)	(0.261)	(0.206)	(0.150)
Instate	0.104	-0.114	-0.187	-0.394
	(0.193)	(0.253)	(0.276)	(0.315)
*Instate*Public*	1.466+	0.437	0.475	1.021+
	(0.383)	(0.418)	(0.347)	(0.369)
N	1953	1311	1905	2420

Note: * p<10% ** p<5% ***p<1%. Standard errors are in parentheses.

Table 4.6: Conditional Logit on College Decision: with Extra Control

		I	II	III	IV
		Most Selective	Highly Selective	Very Selective	Selective
Independent variable: 1=Accepted; 0=unaccepted					
Female		0.124	0.0749	-0.113	0.0830
		(0.167)	(0.290)	(0.243)	(0.243)
Race					
Asian		0.0748	-0.0394	0.0452	0.332
		(0.185)	(0.436)	(0.397)	(0.366)
Black		1.205+	0.111	0.0275	0.0191
		(0.435)	(0.600)	(0.639)	(0.345)
Hispanic		0.919**	0.525	0.473	0.823*
		(0.359)	(0.564)	(0.502)	(0.435)
Other		0.615	-0.375	-0.0589	0.269
		(0.383)	(0.618)	(0.510)	(0.528)
Family Income					
	2	-0.599*	-0.125	-0.367	0.203
		(0.358)	(0.704)	(0.507)	(0.356)
	3	-0.184	-0.453	-0.0784	0.0895
		(0.343)	(0.681)	(0.477)	(0.373)
	4	-0.218	-0.406	0.118	0.246
		(0.360)	(0.694)	(0.500)	(0.430)
Mother's Education					
	2	0.251	0.197	0.316	0.0605
		(0.288)	(0.483)	(0.346)	(0.270)
	3	0.0450	0.00560	0.201	0.299
		(0.261)	(0.402)	(0.342)	(0.302)
SAT Math		0.00371+	0.00471**	0.00378**	0.00269
		(0.00135)	(0.00215)	(0.00167)	(0.00165)
SAT Verbal		0.00183	0.00358*	0.00149	0.000826
		(0.00114)	(0.00203)	(0.00155)	(0.00146)
GPA		0.587**	0.202	0.429	0.0132
		(0.253)	(0.343)	(0.299)	(0.241)
Taking AP course		-0.0309	-0.175	0.298	-0.0467
Extra curricular					
		(0.229)	(0.304)	(0.296)	(0.332)
Sports		0.307*	0.00603	-0.00730	0.00206
		(0.180)	(0.306)	(0.274)	(0.264)
Music		0.139	0.195	-0.329	-0.0735
		(0.155)	(0.291)	(0.260)	(0.250)
Academic		0.173	-0.156	0.165	0.0463
		(0.271)	(0.348)	(0.252)	(0.231)
Other		0.0692	0.0799	-0.0595	-0.158
		(0.187)	(0.276)	(0.237)	(0.231)
Hours		-0.00928	0.0138	-0.0427	0.00698
		(0.0517)	(0.0964)	(0.0810)	(0.0743)
Volunteer		-0.176	0.111	-0.213	0.0628
		(0.316)	(0.452)	(0.331)	(0.281)
Instate		0.123	0.837*	-0.0514	-0.118
		(0.247)	(0.434)	(0.444)	(0.560)
*Instate*Public*		1.897+	-0.0634	0.578	0.446
		(0.486)	(0.728)	(0.558)	(0.648)

For extra controls, see next page

Table 6 cont'

	I Most Selective	II Highly Selective	III Very Selective	IV Selective
Application Information				
Most Selective:				
Dummy Applied		-0.385	0.214	0.384
		(0.372)	(0.395)	(0.549)
Number Applied	0.123+	0.0834	-0.126	-0.279
	(0.0339)	(0.118)	(0.171)	(0.311)
Acceptance Rate	7.401+	0.685	0.178	0.331
	(0.413)	(0.558)	(0.690)	(0.951)
Highly Selective:				
Dummy Applied	-0.874**		0.0660	-0.293
	(0.364)		(0.394)	(0.522)
Number Applied	0.00395	-0.291+	-0.176	0.218
	(0.0856)	(0.112)	(0.163)	(0.360)
Acceptance Rate	0.649**	9.667+	0.305	0.354
	(0.309)	(0.769)	(0.401)	(0.516)
Very Selective:				
Dummy Applied	-0.104	0.654		-0.109
	(0.494)	(0.577)		(0.365)
Number Applied	-0.111	-0.0439	-0.118	0.0828
	(0.127)	(0.115)	(0.0996)	(0.164)
Acceptance Rate	-0.124	-0.501	10.64+	-0.0432
	(0.456)	(0.492)	(0.742)	(0.387)
Selective:				
Dummy Applied	-0.337	-0.660	-0.731	
	(0.693)	(0.824)	(0.625)	
Number Applied	-0.0801	-0.0216	0.157	-0.232**
	(0.193)	(0.247)	(0.113)	(0.0934)
Acceptance Rate	0.553	0.551	0.224	10.93+
	(0.654)	(0.692)	(0.571)	(0.693)
Not Selective:				
Dummy Applied	-1.540	-0.412	-1.144	-0.374
	(1.369)	(1.470)	(1.763)	(0.597)
Number Applied	0.696	0.0654	1.590	-0.0953
	(0.937)	(0.848)	(1.384)	(0.274)
Acceptance Rate	0.932	0.467	-0.261	0.393
	(0.995)	(1.317)	(0.876)	(0.495)
N	1953	1311	1905	2420

Note: * p<10% ** p<5% ***p<1%. Standard errors are in parentheses.

As stated in model part, one thing that threatens the validity of conditional logit is the violation of IIA assumption. For our estimation, the issue boils down to a question whether the addition of some students may change the probability of admitting other students with or without similar characteristics of new alternatives. Therefore, I run two sets of additional regressions, which randomly drop half of applicants or drop all white students, and compare the results with original regressions. In all settings, the P value of Hausman tests are greater than 0.5.

Tables 5 shows factors that matter when colleges are making admission decisions. Minority groups have certain advantages when applying to most selective and highly selective colleges. Other things being equal, students in the lowest family income category are more likely to be admitted by most selective colleges. Both SAT math and verbal score increases one's possibility to be admitted for colleges in all selectivity level. As for academic performance in high school, higher GPA increases one's probability to be admitted and attending AP class increase one's probability to be admitted by most selective colleges.

Table 6 added information generated from one's college application process. For the admission of most selective college, clearly the number of application and acceptance rate for own category has significantly positive impact on the probability of admission. The more applications she files, the more likely shes accepted by one college. Since the college does not observed the number of colleges shes applied to, it means the increase in the number of application not only serves to increase ones chances to be admitted by at least one selective college, but also indicates that this individual has higher ability unobservable to econometricians but observable to colleges and students such that she got better chances to be admitted. For the same set of students and their application to colleges in the second category (Highly selective colleges), although an increase in the acceptance rate by colleges in the second category (which indicates higher ability) implies higher chances to be admitted by a most selective college, the behavior of applying to second category college decreases ones chances to be admitted by a first category college. This result is consistent with our expectation about safety school: the more safety school one applies, the lower her evaluation on her ability and thus lower chance to be admitted by a top college. Also it's consistent with the number of applications shown in table 4 and figure 3: compared to those who do

not apply to safety school, those who apply to safety school may have lower ability and lower chances to be admitted.

Even we assume that colleges could learn about applications to other colleges using FAFSA information, colleges do not know the admission decisions made by other colleges. Therefore the huge impact on acceptance rate by other colleges confirms our supposition that the admission result could indicate students' ability.

In the rest columns, the dummy that applied to a college in higher selectivity level is always positive and the dummy that applied to a college in lower selectivity level is always negative. However, since theyre not significant, we cannot conclude that firmly. Another interesting finding is that the number of applications filed to own selectivity level. For all columns the coefficients are significantly negative except for most selective college. This may be understandable: the intensity of competition in the most selective level is very high while that in the rest categories are relatively low. Therefore, the increase in the number of applications to most selective colleges indicates higher personal aspiration but the increase in the number of applications to other colleges indicates lower self-evaluation.

The addition of application information significantly changed the coefficients on gender and race dummies. Female dummy is significantly negative for the two highest selectivity level in table 5 while the effect disappeared in table 6. Table 5 implies that females are discriminated against but the result in table 6 may suggest that systematically males present higher ability or better performance which can not be captured by standardized indicators. For example, the stylized fact is that females usually have higher GPA than men regardless of their similar SAT performance (see Conger, 2015 for a summary). On the other hand, however, it may also suggest a discrimination story: women tend to under-perform in application as they know their chances are low.

The addition of application information also significantly decreased the advantage of minorities. At the first sight, it may imply the uncontrolled model overestimated the preferential treatment to ethnic minorities. However, the application behavior may also capture a part of preferential treatment, that is, one is more likely to apply when she knows shes pretty likely to be admitted. For example, Black & Cortes (2015) run conditional logit model (based on student fixed effect) separately for different ethnic groups. Using a large

database, they found that for all ability groups, black students are more likely to file applications to better colleges, i.e. those with higher mean SAT score than their own, than Hispanics than white than Asian students. Therefore, The extra variables about application does not capture unobserved advantage of being admitted from an objective way, but rather the unobserved advantage being admitted from a perceived way.

As the dataset ELS2002 contains information about FAFSA, I could further test whether it affected colleges decision. In the dataset, attempts were made to match respondents to records in the U.S. Department of Education Central Processing System (CPS), which contains detailed data collected from FAFSA. The restricted ELS2002 dataset provides full FAFSA records in several years starting from 2004-2005. Since most students (91%) were enrolled to college in 2004-2005 academic year, I use the 2004-2005 records since its the one that colleges refer to when they make admission decisions. Each record lists all SIX colleges the student wants to send application to using the original ORDER listed by the student.

The attempt to matching respondents to FAFSA record was not applied to all respondents. The following table shows the attempt/matching results for ELS2002 full sample and our sample who applied to 4yr colleges only.

Table 4.7: Matching ELS2002 to FAFSA data

	ELS2002 full sample	Only applied to 4yr
No attempt	3088	1068
Not matched	6625	1725
Matched	6484	3857
Total	16197	6650

With the existence of no attempt cases, it is hard to analyze all students who only applied to 4yr colleges. Therefore I limit the following analysis to those who were sure to have applied to FAFSA, i.e. 3857 respondents who were matched to FAFSA record. Also, to exclude those who entered college earlier or later (whose admission decision was not affected by preference revealed by 04/05 FAFSA), I only kept first year student in 04/05, which left 3775 students. Then I rerun regression about colleges' admission decisions using similar settings as is shown in table 6, adding dummies showing that whether the student listed the college at 1st, 2nd, 3rd, 4th, 5th, 6th notch or not .Table 8 shows regression result (since coefficients on other variables are similar to that in table 6, I only list coefficients of FAFSA

college preferential orders). The six dummies indicating the preference order of the college

Table 4.8: Conditional Logit on College Decision: with FAFSA information

	I	II	III	IV
	Most Selective	Highly Selective	Very Selective	Selective
Independent variable: 1=Accepted; 0=unaccepted				
(Other variables as in table 6 omitted)				
...				
FAFSA preferential orders:				
1st	0.927+	-1.203*	3.563+	3.532+
	(0.323)	(0.683)	(0.779)	(0.645)
2nd	0.630*	0.754	0.501	1.911+
	(0.335)	(0.637)	(0.626)	(0.597)
3rd	0.693*	0.0382	1.705**	3.736+
	(0.356)	(0.763)	(0.851)	(0.883)
4th	0.802**	1.774**	1.828**	-0.268
	(0.352)	(0.87)	(0.835)	(0.729)
5th	1.311+	1.175	1.674	1.642
	(0.38)	(1.086)	(1.067)	(1.077)
6th	0.78	1.549	-0.141	-0.906
	(0.524)	(1.171)	(1.315)	(1.143)
N	1183	766	1021	1421

Note: * p<10% ** p<5% ***p<1%. Standard errors are in parentheses.

are significant, compared to not sending FAFSA at all. It means that other things equal, colleges are more/less likely to admit those who sent FAFSA application but the order of preferences shown in the list shows no systematic patterns. This result suggests that FAFSA variables affect admission result mainly because students send FAFSA to colleges they have more confidence in being admitted rather than colleges use it to obtain students' preference; otherwise, coefficient on being the first college should be greater than being the second college than being the third, etc.

4.5 College Matriculation

This section focuses on student's matriculation decision given several colleges have accepted the student. Previous literature suffers from the lack of detailed data on the exact choice set of an applicant, i.e. all institutions that have accepted the applicant after the second stage. As was discussed by Klaauw (2002), when the choice sets containing all alternatives

of an applicant cannot be observed, it becomes very difficult to distinguish the impact of a college's own characteristics from that of other colleges. Then omitted variable bias arises since the related variables cannot be considered exogenous with respect to the matriculation decision.

4.5.1 Methodology

The ELS2002 data has a huge advantage that it contains rich records on college application and admission process and therefore detailed information about an applicants full choice set in the final stage is identified. At the beginning of this stage, assuming the student has several colleges that have accepted her and she will choose the most desirable one to matriculate. Suppose each alternative could be characterized by a vector of Y that includes characteristic of the college including tuition, financial aid, and reputation etc. A students individual-specific preference on the college also depends on the matching quality between her and the college, denoted as Z, which includes the distance between the college and her hometown, the difference between her academic performance and that of the average student body at the college etc. Apart from those college-specific characteristics, the students personal characteristics X, like her past academic performance and family background may also affect her preference on different colleges. Therefore in a generalized model we assuming the utility of individual i choosing college j is a linear function of those variables:

$$U_{ij} = X_i\alpha_j + Y_j\beta + Z'_{ij}\gamma + \epsilon_i$$

Since she is faced with a limited number of choice, the random utility model (Mcfadden, 1974) could be used by assuming that individual i will choose college J if and only if her utility from J is greater than her utility from other colleges in the choice set.

$$D_{iJ} = 1 \; if \; U_{iJ} \geq U_{ij} \; for \; all \; j \neq J$$

Therefore, the probability of choosing school j given the choice set and all related variables is

$$Prob(Y_i = j) = \frac{exp(x_i\beta_j + z'_{ij}\gamma)}{\sum_{j=1}^{J}(x_i\beta_j + z'_{ij}\gamma)}$$

In this process, a problem for conditional logit is that the independence from irrelevant alternative (IIA) does not necessarily hold. Thus, we also add a nested conditional logit, assuming that colleges of the same characteristics are substitutable.

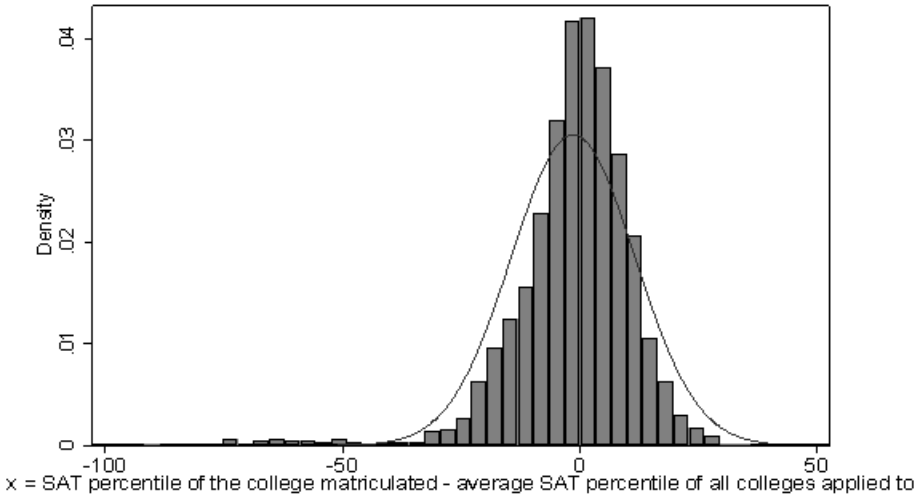
4.5.2 Empirical Results

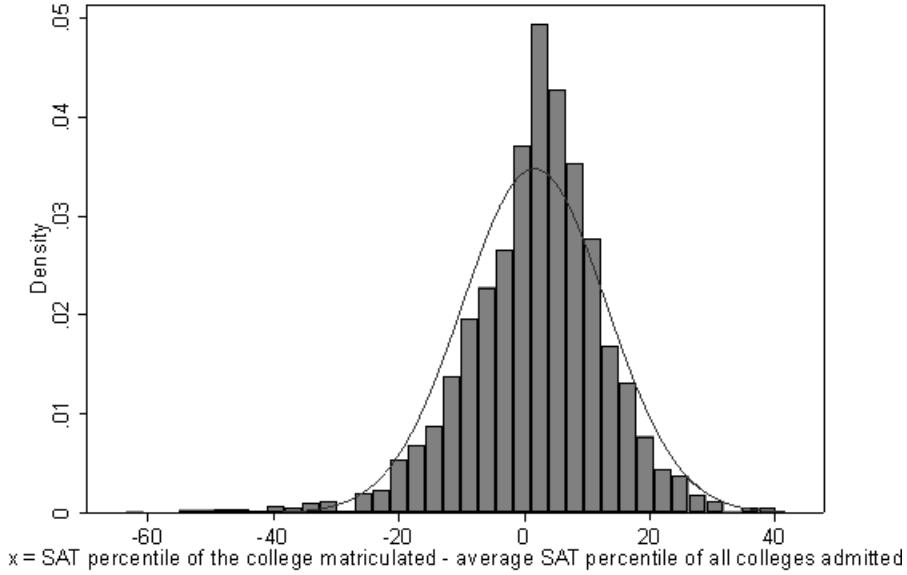
In this step, the record shows that there are 12,471 applications eventually accepted by colleges. For all applications with no missing variables, the dataset contains 91 individuals who were accepted by only one college, and 3130 individuals who have been accepted by at least two colleges. Since both Conditional Logit and Nested Logit required multiple observations for each group, i.e. individual student, the regression sample only include those accepted by at least two colleges. This leaves 10149 individual-choice paired observations. The average number of acceptance per student is 3.2.

This section focuses on the matriculation decision. If we assume that students have an estimation about their own ability and apply to colleges accordingly, there should be a relationship between all colleges he applied to / admitted by and the college he eventually matriculate to. Figure 4 shows the distribution of the difference between SAT percentile of the student eventually matriculated and that of average college he applied to/admitted by. For both distribution, the peaks occur when SAT percentile of matriculated college are the same as the average SAT percentile of students' application or admission. It shows that students tend to go to a college that matches her ability estimation. At the same time, they also like the college slightly above their estimation than that slightly below their estimation. Therefore, I choose to add two variables in matriculation decision: the absolute value of average student body's SAT percentile of the college he was admitted minus students' SAT percentile, and whether the difference is positive or negative.

In regressions, the dependent variable is whether one has matriculated into the college or not. As a student can at most matriculate into one college, only one of those multiple

Figure 4.4: Distribution of SAT difference





observations corresponding to a specific individual has dependent variable equal to 1; those who did not matriculate into any college are dropped in conditional logit and nested logit model.

I first used conditional logit model to estimate factors affecting matriculation decision and a test is required to examine whether IIA condition holds. As is suggested by Hausman and McFadden (1984), if a subset of alternatives is truly irrelevant with other alternatives, omitting that subset in the regression will not lead to inconsistent estimates. Therefore, Hausmans test can be applied. Following this spirit, I rerun the regression five times after dropping each selectivity categorie. For all comparisons between full sample and sub-sample regressions, Hausman test rejects IIA property at 1% level of significance. Therefore, the nested logit model, instead of conditional logit model, is supposed to produce unbiased estimations.

Column I reports conditional logit result. Column II reports nested logit result. In Column III, the interaction between selectivity-level-specific coefficients and individual characteristics are added. Since most coefficients are not significant I do not report them in the table.

Though IIA does not hold, the estimated coefficients in Column I and II are pretty similar. The estimated coefficients on financial variables are as expected. An $1,000 increase in tuition will significantly decrease ones probability to matriculate into the college, while financial aid factors like grant, student loan and student work opportunities that reduces the students net payment provide the student more incentives to matriculate. The matriculation probability increases with enrollment size but public or private does not have significant impact on matriculation decision.

The colleges selectivity level significantly affected ones probability to matriculate. The largest impact is exerted when the college is Most Selective. It increased ones probability to matriculate by 170% in odd ratio. Sequentially Highly Selective is more attractive than Very Selective than Selective. However, after adding in individual interaction with level of selectivity, the standard error increased drastically, making the coefficient on selective college not significant.

The other matching variable that matter is college-hometown distance. Living within

Table 4.9: Matriculation Decision

	I Conditional Logit	II Nested Logit	III Nested Logit
Independent variable: 1=Matriculated; 0=Otherwise			
Tuition	-0.0219+	-0.0188+	-0.0165**
	(0.00684)	(0.00690)	(0.00713)
Enrollment	0.0476**	0.0478**	0.0567+
	(0.0211)	(0.0208)	(0.0218)
Public	-0.137	-0.0989	-0.0916
	(0.143)	(0.140)	(0.144)
Barrons Selectivity			
Most Selective	1.695+	1.638+	1.171
	(0.186)	(0.202)	(2.155)
Highly Selective	0.636+	0.622+	0.323
	(0.163)	(0.175)	(1.750)
Very Selective	0.333**	0.410+	1.226
	(0.147)	(0.157)	(1.485)
Selective	0.0666	0.134	0.633
	(0.135)	(0.148)	(1.283)
Instate	-0.0557	-0.0173	-0.0370
	(0.0967)	(0.0945)	(0.0978)
Grant	0.973+	0.922+	0.924+
	(0.0906)	(0.0970)	(0.101)
Loan	1.593+	1.555+	1.624+
	(0.116)	(0.124)	(0.131)
Student work	1.282+	1.276+	1.219+
	(0.134)	(0.143)	(0.145)
Distance_community	0.182**	0.180**	0.185**
	(0.0765)	(0.0735)	(0.0769)
Distance_faraway	-0.116	-0.129	-0.143*
	(0.0848)	(0.0823)	(0.0847)
Sat difference	-0.0116+	-0.00974+	-0.00768*
	(0.00304)	(0.00300)	(0.00402)
Sat difference_plus	0.175	0.116	0.0778
	(0.154)	(0.152)	(0.156)
Interaction with individual variables	no	no	yes
N	7865	7600	7232

Note: * p<10% ** p<5% ***p<1%. Standard errors are in parentheses.

50 miles of the college make students much likely to matriculate and more than 200 miles away from the college decreased students probability to matriculate.

Difference in student-college SAT percentile also matters. A colleges average SAT score of newly enrolled students is a good proxy of academic ability of student body and students care about matching qualities. The result shows that after controlling college selectivity, the greater the difference between students SAT score and the average of the college, the lower the chance the student will matriculate. Though the indicator for "positive difference" has positive impact, it's not significant.

4.6 Conclusion

Students' transition from high school to college is a multi-stage process that involves repeated interaction between students and colleges. Due to paucity of data, previous research usually sees the whole process as one single decision made by students, omitting constraints or limited choice set faced by them. This paper makes effort to disintegrate the process into several stages and examines from the view of decision makers what factors affect college admission and matriculation results. As for admission decision, we use variables generated from application process to instrument student ability and obtain unbiased estimation. As for matriculation decision, with complete information about students' choice set, we also obtained unbiased estimation.

References

[1] Avery, C., and Hoxby, C. M. Do and should financial aid packages affect students' college choices? In *College choices: The economics of where to go, when to go, and how to pay for it*. University of Chicago Press, 2004, pp. 239–302.

[2] Babcock, P., Bedard, K., and Schulte, J. No cohort left behind? *Journal of Urban Economics 71*, 3 (2012), 347–354.

[3] Barro, R. J., and Lee, J.-W. Sources of economic growth. In *Carnegie-Rochester conference series on public policy* (1994), vol. 40, Elsevier, pp. 1–46.

[4] Behrman, J. R., and Birdsall, N. The reward for good timing: Cohort effects and earnings functions for brazilian males. *The Review of Economics and Statistics* (1988), 129–135.

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

[6] Black, S. E., Cortes, K. E., and Lincove, J. A. Apply yourself: Racial and ethnic differences in college application. Tech. rep., National Bureau of Economic Research, 2015.

[7] Bound, J., and Turner, S. Cohort crowding: How resources affect collegiate attainment. *Journal of public Economics 91*, 5 (2007), 877–899.

[8] Card, D., and Krueger, A. B. Labor market effects of school quality: Theory and evidence. Tech. rep., National Bureau of Economic Research, 1996.

[9] Card, D., and Lemieux, T. Dropout and enrollment trends in the postwar period. *Risky Behavior among Youths: An Economic Analysis* (2009), 439.

[10] Chandra, N. K. Education in china: From the cultural revolution to four modernisations. *Economic and Political Weekly* (1987), AN121–AN136.

[11] Chen, H.-j. Educational systems, growth and income distribution: a quantitative study. *Journal of Development Economics 76*, 2 (2005), 325–353.

[12] Chen, S. Fertility of women during the 42 year period from 1940 to 1981.

[13] Conger, D. High school grades, admissions policies, and the gender gap in college enrollment. *Economics of Education Review 46* (2015), 144–147.

[14] Connelly, R. A framework for analyzing the impact of cohort size on education and labor earning. *Journal of Human Resources* (1986), 543–562.

[15] Coxhead, I., Shen, G., and Yao, Y. Do schools cause students? evidence from chinas post-famine baby boom.

[16] Dale, S. B., and Krueger, A. B. Estimating the Payoff to Attending a More Selective College: An Application of Selection on Observables and Unobservables. *The Quarterly Journal of Economics 117*, 4 (Nov. 2002), 1491–1527.

[17] Dale, S. B., and Krueger, A. B. Estimating the Effects of College Characteristics over the Career Using Administrative Earnings Data. *Journal of Human Resources 49*, 2 (Mar. 2014), 323–358.

[18] Dillon, E. W., and Smith, J. A. Determinants of the match between student ability and college quality. *Journal of Labor Economics 35*, 1 (2017), 45–66.

[19] Easterlin, R. A. What will 1984 be like? socioeconomic implications of recent twists in age structure. *Demography 15*, 4 (1978), 397–432.

[20] Easterlin, R. A., Wachter, M. L., and Wachter, S. M. Demographic influences on economic stability: The united states experience. *Population and development review* (1978), 1–22.

[21] Epple, D., Romano, R., and Sieg, H. Peer effects, financial aid and selection of students into colleges and universities: An empirical analysis. *Journal of Applied Econometrics 18*, 5 (2003), 501–525.

[22] Fang, H., Eggleston, K. N., Rizzo, J. A., Rozelle, S., and Zeckhauser, R. J. The returns to education in china: Evidence from the 1986 compulsory education law. *NBER Working Paper Series* (2012), 18189.

[23] Fertig, M., Schmidt, C. M., and Sinning, M. G. The impact of demographic change on human capital accumulation. *Labour Economics 16*, 6 (2009), 659–668.

[24] Fleisher, B. M., Li, H., Shi, L., and Wang, X. Sorting, selection, and transformation of the return to college education in china.

[25] Fleisher, B. M., and Wang, X. Returns to schooling in china under planning and reform. *Journal of Comparative Economics 33*, 2 (2005), 265–277.

[26] Freeman, R. B. The effect of demographic factors on age-earnings profiles, 1979.

[27] Glewwe, P., and Kremer, M. Schools, teachers, and education outcomes in developing countries. *Handbook of the Economics of Education 2* (2006), 945–1017.

[28] Graber, G. Using FAFSA data ethically and strategically for enrollment management, Jan. 2014. 00000.

[29] Guo, R., Yi, J., and Zhang, J. Family size, birth order, and tests of the quantity–quality model. *Journal of Comparative Economics* (2016).

[30] Han, M., and Yang, X. Educational assessment in china: Lessons from history and future prospects. *Assessment in Education: principles, policy & practice 8*, 1 (2001), 5–10.

[31] Hansmann, H. Nonprofit enterprise in the performing arts. *The Bell Journal of Economics* (1981), 341–361.

[32] Hanushek, E. A. The trade-off between child quantity and quality. *Journal of political economy* (1992), 84–117.

[33] Hanushek, E. A., and Kimko, D. D. Schooling, labor-force quality, and the growth of nations. *American economic review* (2000), 1184–1208.

[34] Harmon, C., and Walker, I. The returns to the quantity and quality of education: evidence for men in england and wales. *Economica 67*, 265 (2000), 19–35.

[35] Hartog, J., Sun, Y., and Ding, X. University rank and bachelor's labour market positions in china. *Economics of Education Review 29*, 6 (2010), 971–979.

[36] Hausman, J., and McFadden, D. Specification tests for the multinomial logit model. *Econometrica: Journal of the Econometric Society* (1984), 1219–1240.

[37] Hongyi, L., and Huang, L. Health, education, and economic growth in china: Empirical findings and implications. *China Economic Review 20*, 3 (2009), 374–387.

[38] Horstschrer, J. University rankings in action? The importance of rankings and an excellence competition for university choice of high-ability students. *Economics of Education Review 31*, 6 (2012), 1162–1176.

[39] Hoxby, C., and Avery, C. The missing" one-offs": The hidden supply of high-achieving, low-income students. *Brookings papers on economic activity 2013*, 1 (2013), 1–65.

[40] Huang, M.-L., and Hsieh, Y.-S. *Journal of Population Studies (Renkou Xuekan)*, 49 (2014), 1–40.

[41] Hurwitz, M. The impact of legacy status on undergraduate admissions at elite colleges and universities. *Economics of Education Review 30*, 3 (2011), 480–492.

[42] James, E. Why do different countries choose a different public-private mix of educational services? *Journal of human resources* (1993), 571–592.

[43] Lee, J.-W., and Barro, R. J. Schooling quality in a cross–section of countries. *Economica 68*, 272 (2001), 465–488.

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

[45] Li, H., Zhang, J., and Zhu, Y. The quantity-quality trade-off of children in a developing country: Identification using chinese twins. *Demography 45*, 1 (2008), 223–243.

[46] Lin, J., and Chen, Q. Academic pressure and impact on students'development in china. *McGill Journal of Education/Revue des sciences de l'éducation de McGill 30*, 002 (1995).

[47] Lincove, J. A., and Cortes, K. E. Match or Mismatch? Automatic Admissions and College Preferences of Low-and High-Income Students. Tech. rep., National Bureau of Economic Research, 2016.

[48] Ling, M. bad students go to vocational schools!: Education, social reproduction and migrant youth in urban china. *China Journal*, 73 (2015), 108–131.

[49] Long, M. C. College quality and early adult outcomes. *Economics of Education review 27*, 5 (2008), 588–602.

[50] Luppino, M., and Sander, R. College major peer effects and attrition from the sciences. *IZA Journal of Labor Economics 4*, 1 (2015), 4.

[51] Ma, Y. Impact of birth cohort size on education achievement of cohort members. *Population Research (Renkou Yanjiu) 38*, 3 (2014), 67–80.

[52] Manski, C. F., and Wise, D. A. *College choice in America*. Harvard University Press, 1983.

[53] Mauger, P. Changing policy and practice in chinese rural education. *The China Quarterly 93* (1983), 138–148.

[54] Meng, X., and Gregory, R. Exploring the impact of interrupted education on earnings: the educational cost of the chinese cultural revolution.

[55] Monks, J., and Ehrenberg, R. G. The Impact ofU. S. News & World Report-College Rankings on Admissions Outcomes and Pricing Policies at Selective Private Institutions.

[56] Nurnberg, P., Schapiro, M., and Zimmerman, D. Students choosing colleges: Understanding the matriculation decision at a highly selective private institution. *Economics of Education Review 31*, 1 (2012), 1–8.

[57] Ping, T. Trends and regional differentials in fertility transition. In *The Changing Population of China*. Blackwell Publishers, 2000.

[58] Reiling, R. B. Does size matter? educational attainment and cohort size. *Journal of Urban Economics 94* (2016), 73–89.

[59] Rosen, S. Recentralization, decentralization, and rationalization: Deng xiaoping's bifurcated educational policy. *Modern China 11*, 3 (1985), 301–346.

[60] Saavedra, J. E. Resource constraints and educational attainment in developing countries: Colombia 1945–2005. *Journal of Development Economics 99*, 1 (2012), 80–91.

[61] Scott, E. Extension of grammar schools and selection in education. Tech. rep., House of Lords Library Notes.

[62] Smith, J., Pender, M., and Howell, J. The full extent of student-college academic undermatch. *Economics of Education Review 32* (2013), 247–261.

[63] Snyder, T. D., and Dillow, S. A. Digest of education statistics, 2012. nces 2014-015. Tech. rep., National Center for Education Statistics, 2013.

[64] Song, Y. Poverty reduction in china: The contribution of popularizing primary education. *China & World Economy 20*, 1 (2012), 105–122.

[65] Tobias, J. L. Model uncertainty and race and gender heterogeneity in the college entry decision. *Economics of Education Review 21*, 3 (2002), 211–219.

[66] Tobias, J. L. The effects of cognitive ability and high school quality on college entry decisions: nonparametric estimation of parameters of interest. *Applied Economics 35*, 2 (2003), 209–215.

[67] Toutkoushian, R. K. Do parental income and educational attainment affect the initial choices of New Hampshire's college-bound students? *Economics of Education Review 20*, 3 (2001), 245–262.

[68] Tsang, M. C. Financial reform of basic education in china. *Economics of Education Review 15*, 4 (1996), 423–444.

[69] Unger, J. *Education under Mao: Class and competition in Canton schools, 1960-1980*. Columbia University Press, 1982.

[70] Van der Klaauw, W. Estimating the effect of financial aid offers on college enrollment: A regressiondiscontinuity approach. *International Economic Review 43*, 4 (2002), 1249–1287.

[71] Wachter, M. L., and Wascher, W. L. Leveling the peaks and troughs in the demographic cycle: an application to school enrollment rates. *The Review of Economics and Statistics* (1984), 208–215.

[72] Weiler, W. C. Factors influencing the matriculation choices of high ability students. *Economics of Education Review 15*, 1 (1996), 23–36.

[73] Welch, F. Effects of cohort size on earnings: the baby boom babies' financial bust. *The Journal of Political Economy* (1979), S65–S97.

[74] Winston, G. C. Subsidies, hierarchy and peers: The awkward economics of higher education. *The Journal of Economic Perspectives 13*, 1 (1999), 13–36.

[75] Wu, X., and Xie, Y. Does the market pay off? earnings returns to education in urban china. *American Sociological Review* (2003), 425–442.

[76] Yang, D. *Equal Education in China: the Ideal and Reality*. Peking University Press, 2006.

[77] Yang, D. Dealing with the school selection in promotion to middle school in beijing. Tech. rep., 21st Century Education Research Institute, 2011.

[78] Yao, Y. Thesis on employment policy and college graduate employment in china. *Population Journal (Renkou Xuekan)*, 4 (2004), 8–11.

[79] Yu, L., and Suen, H. K. Historical and contemporary exam-driven education fever in china. *KEDI Journal of Educational Policy 2*, 1 (2005).

[80] Zhang, J. Socioeconomic determinants of fertility in china. *Journal of Population Economics 3*, 2 (1990), 105–123.

[81] Zhao, L. Enrollment quota control, elite selection and access to education in rural china. *International Education Studies 3*, 1 (2010), 23.

[82] Zhou, L. The incentive and cooperation of government officials in the political tournaments: An interpretation of the prolonged local protectionism and duplicative investments in china. *Economic Research Journal (Jingji Yanjiu) 6* (2004), 33–40.

[83] Zhou, L. Governing china's local officials: An analysis of promotion tournament model. *Economic Research Journal (Jingji Yanjiu) 7* (2007), 36–50.

[84] Zhou, X. Economic transformation and income inequality in urban china: evidence from panel data. *American Journal of Sociology* (2000), 1135–1174.