The Dissertation Committee for Ziwei He Certifies that this is the approved version of the following dissertation:

Essays on Economics of Education

Committee:

Leigh L. Linden, Supervisor

Richard J. Murphy

Kate Antonovics

Michael Geruso

Essays on Economics of Education

by

Ziwei He, B.S. M.S.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

May 2017

Acknowledgements

I would like to thank Leigh L. Linden, Richard Murphy, Sandra E. Black, Mike Geruso, Kate Antonovics, Julie B. Cullen and Stephen J. Trejo for their invaluable guidance. I would also like to thank the Texas Education Resource Center for providing data. Lastly, I'd like to acknowledge my father's support for my graduate study.

Essays on Economics of Education

Ziwei He, Ph.D

The University of Texas at Austin, 2017

Supervisor: Leigh L. Linden

This dissertation investigates the relationship between various input for education (such as class structure and academic environment) and academic outcomes (including majors and test scores). It also looks at the impact of education on individuals' non-academic outcomes. This dissertation consists of three essays. The first chapter examines the impact of assigning students into different tracks on students' academic performance and subject specialization in China. I make use of regression discontinuity design and find that track assignment significantly affects choice of majors and test scores in high school. For students around the tracking threshold, being assigned to a high track reduces the probability of choosing the science major by 7 percent for boys and 21 percent for girls. The second chapter examines the impact of international peers on domestic students STEM degree in U.S.. I use historical enrollment patterns as an instrumental variable to predict current enrollment of international students and find that the composition and ability of international peers significantly affect the likelihood of graduating with a STEM degree for female and minority domestic students. The third chapter explores the casual relationship from education to religious beliefs in China. I exploit the change in compulsory school law in China in 1986 and find that one additional year of schooling reduce the probability of being religious by 8 percent.

Table of Contents

List of Tables

List of Figures

Introduction

This dissertation examines two questions. First, it explores how students' academic outcomes are affected by educational inputs such as teachers' quality, parental investment and peer effect[1]. Second, it examines the impact of education on religious beliefs. The first chapter analyzes how educational tracking affects test scores and students' choice of high school majors. This work informs the heated debate on the extent to which students should be tracked into different classes based on previous academic performance. Although parents and students frequently spend considerable efforts and resources to secure a seat in the higher track, it remains unclear whether the higher track leads to improvements in the performance of the marginal attendants. In particular, those at the bottom of test score distribution in the higher track may be adversely affected by the more demanding and competitive environment.

To understand the impact of track assignment on the marginal attendants, I collect data from a Chinese high school and use a regression discontinuity approach where assignment into high-ability classes relies solely on school entrance exam scores. I find that being assigned to the high track significantly reduces the likelihood of choosing the more challenging science major for both genders. This is the first paper to show that being assigned to higher track discourages students in the bottom of score distribution

[1]This dissertation utilizes confidential data from the State of Texas supplied by the Texas Education Research Center (ERC) at The University of Texas at Austin. The author gratefully acknowledges the use of these data. The conclusions of this research do not necessarily reflect the opinion or official position of the Texas Education Research Center, the Texas Education Agency, the Texas Higher Education Coordinating Board, the Texas Workforce Commission, or the State of Texas. Any errors are attributable to the author.

from majoring in science even though it has non-negative effect on students' academic performance. I further explore the possible mechanisms to explain why marginal students in the high track are less likely to major in science. My findings suggest that peers and relative abilities play important roles in the choices of majors.

To further explore the impact of peer ability on major choice in college, my second chapter examines the effects of international students on native students' choice of Science Technology Engineering and Math (STEM) majors in college. I use two measures for the quality of international peers: first, I identify whether each individual student's sending country is English-speaking or not. Second, I use the test scores from the Programme for International Student Assessment and calculate the average peer scores for each school. To address the endogeneity of admission of international students, I construct an instrumental variable that allocates recent flows based on prior enrollment patterns. I find that an increasing quality of international peers decreases the number of U.S.-born college students who graduated with degrees in STEM fields. My findings also suggest that, compared to male students, female students are more likely to leave STEM education.

My third chapter examines the impact of education on religious beliefs. It is the first study to explore this topic in a developing country. We make use of the compulsory school law implemented in China in 1986 that extended schooling from 6 to 9 years. We find that individuals who experienced the reform had more years of education, which led to lower levels of religious belief and less participation in religious activities later in life.

Chapter 1. The Impact of Academic Tracking Assignment on Students' Subject Specializations

1.1 INTRODUCTION

Education is an important determinant of many life outcomes, including wages and health. How to improve the efficiency of educational systems is one of the central topics in public policy debate internationally. This is particularly true in developing countries where resources are limited. One controversial topic in this discussion is the extent to which students should be tracked into different classes based on ability as measured by previous academic performance[2]. Students in higher tracks may benefit from better peers, teachers and resources. While parents and students frequently expend considerable effort and resources to secure a seat in a higher track, it remains unclear whether the higher track leads to improvements in the performance of the marginal attendant since those who are in the bottom of test score distribution in the higher track may experience more stress caused by more demanding curriculums and competitive environment. Understanding the impact of track assignment not only helps to settle this controversy but also has important educational policy implications.

The general problem researchers have encountered is that students who choose the high track are inherently different from these who choose the low track. To avoid this selection problem, several papers make use of the tracking rule and use regression

[2] Students could be tracked into different schools at the same time. This paper only examines within school tracking.

discontinuity design to exploit quasi-random tracking assignment (Duflo et al 2011; Vardardottir 2013; Ma and Shi 2014). Identification relied on the assumption that students around the tracking threshold are similar in abilities and other unobservable characteristics although they are assigned to different tracks. This paper adopts the same identification strategy but better experimental setting in which students are less likely to be able to manipulate the threshold[3]. I collect data on five cohorts (2008-2012) of 10th grade students in a Chinese high school, where students are tracked solely based on performance on an entrance exam. This generates a discontinuity, students who score above a critical score are placed in the high track and those below in the low track. There is little room for manipulation since the test are graded anonymously, tracking threshold changes year to year[4], and it is unknown to students and graders. I compare students in the same school who attend different tracks because their test scores are just above or below the threshold for tracking.

The main outcome of interest is the effect of being assigned to high track on choice of major and academic achievement in high school, comparing students around the threshold who are otherwise assumed to be identical. At the end of 10th grade in China, one year after tracking, students are required to declare a broad major (science or social science) that they intend to study in college[5]. These majors have different high school curriculums, exams and College Entrance Exams. Therefore, a student's choice of

---

3 Students are less like to manipulate the tracking threshold since it is unknown to public and changes year to year.

4 Since the high school entrance exam is not comparable across years, the tracking thresholds are not the same.

5 Since the educational reform in 2014, students do not choose majors in high school. Source: http://www.gov.cn/zhengce/content/2014-09/04/content_9065.htm

major in high school has direct implications for the choice set of majors in college. For example, majoring in science in high school is a pre-requisite for choosing STEM majors at the college level. I find that although in general students with higher test scores are more likely to choose science majors, students who are just above the threshold and enrolled in the high track are 21 and 7 percentage points less likely to major in science for girls and boys respectively.

The second outcome of interest is the impact of admittance to the high track on academic achievement, which is measured by an exam one semester after being tracked. This exam is employed because it is the same for both tracks and takes place before students choose majors. I find that tracking increase the exam scores of the marginal male students assigned to high track by 0.16 standard deviations while it has an insignificant effect on the test score of girls. This mask large heterogeneity across subjects: boys increase math, science and social science, while girls improve Chinese, English and social science but experience no significant improvement in overall grades.

This paper contributes to the literature by providing a new dimension for understanding the impact of track assignment. Although previous research has only found positive effect of being assigned to high track on academic performance (Vardardottir 2013 and Ma and Shi 2014), my findings suggest that track assignment discourages students from majoring in science. To the best of my knowledge, this is the first paper studying the impact of track assignment on the choice of major[6]. This paper is also

[6] The only other paper examines the choice of major in tracking literature is Dee and Lan (2015). Instead of within-school tracking, they examine across-school tracking and find insignificant impact of tracking on choice of major.

broadly related to the mismatch literature. Arcidiacono *et. al.* (2016) find that minority students would be more likely to graduate with a science degree had they attended a lower ranked university. The lowest scoring students in this paper are similar to these minority students in that they are both at the lower end of the test score distributions among their immediate academic peers. My finding confirms that these students are less likely to major in science in a more competitive environment, despite being exposed to better educational resources.

This paper also complements the existing work on the academic achievement effects of track assignment. Vardardottir (2013) examined the tracking assignment for high school students in Iceland and found that being assigned to a high ability class increases academic achievement by 0.32 standard deviations. Ma and Shi (2014) studied the similar question using a high school in China and concluded that students in high tracking increased their test scores by 0.647 standard deviations. The results in this paper are more comparable to the findings in Ma and Shi (2014) since we obtain data from the same province in China. However, the schools under investigation are very different. First, the high track classes in their dataset have smaller class size and better teachers than low track, while in my setting both tracks have similar class sizes and around fifty percent of teachers teach both tracks. Second, students in Ma and Shi (2014) are reassigned to tracks based on their test scores in each school year, while students would stay in the same track for entire high school in my data. Overall, students in the high track in their paper have better educational inputs than the counterparts in my paper. Therefore, I expect to obtain a smaller estimate compared to the result in their paper.

I also contribute to the important literature on gender differences in education (Fryer and Levitt 2010; Ellison and Swanson 2010; Hyde et al. 2008). I find that the impact on boys is significantly different from the impact on girls. Compared to girls, boys in the high track are more likely to improve their test scores and major in science. Although researchers have shown that educational inputs such as teacher quality and class environment have heterogeneous impact (Ho and Kelman 2014; Carrell *et al.* 2010; Dee 2007 and Gneezy *et al*. 2003), to the best of my knowledge, this is the first paper that suggests track assignment exerts heterogeneous impacts by gender.

I further explore the possible mechanisms why marginal students in the high track are less likely to major in science, including teacher resources, relative abilities, and peer abilities. I find that compared to boys, girls in the high track are less likely to have the same-gender teachers who may serve as better role models for them (Lim and Meer 2015; Lavy and Sand 2015). Those girls are also more likely to be affected by their relative rank and peers.

This paper is related to other recent studies which have examined the impact of relative ability on educational outcomes (Murphy and Weinhardt 2014; Elsner and Isphording 2015(a)(b)). They have used the variation of the test score distribution to obtain the variation in rank and find that a student's ordinal rank improves educational outcomes such as college completion, test scores and task specific confidence, and decrease the probability of engaging in risky behaviors. Although I am not able to identify the mechanisms of major-choice effect, track assignment changes relative ranking for each class and ordinal rank may be one potential channel given the findings in literature.

The rest of the paper is organized as follows: The next section introduces the institutional background. Section 1.3 presents the empirical strategy and the data. Section 1.4 shows the verification of the internal validity of the research design. Section 1.5 presents the estimation results. Section 1.6 explores possible mechanisms, and Section 1.7 concludes.

1.2 Background Information

1.2.1 China's Secondary Education System

In China, primary and secondary education takes 12 years to complete, consisting of primary (1^{st} - 6^{th} grade), junior secondary (7^{th} -9^{th} grade) and senior secondary (10^{th} - 12^{th} grade) stages. Within each province, senior secondary schools (high schools, 10^{th} - 12^{th} grade) are classified by ranks. The ranking is common knowledge to students. Entrance examination scores are the most important determinant of admission to both high schools and colleges[7].

At the end of the 9^{th} grade, students take the High School Entrance Exam (HSEE), a standardized, province-wide test that consists of five parts: Chinese, math, English, social science and science. The science portion includes physics, chemistry and biology while social science includes history, politics and geography. With the aim of accepting a

[7] Students with low tests scores might be able to go to the high track if their parents have a strong relationship with the school.

certain amount of students, high schools set a minimum required score after the exam such that students above the score are admitted with certainty. (The size of the incoming class may vary depending on school size.)

At the beginning of 11th grade, students choose to major in either science or social science. These majors have different high school curriculums. The College Entrance Exam (CEE) takes place at the end of 12th grade and is given differently for two majors. The Chinese and English sections are the same for both majors, while the math section is harder for science majors than for social science majors. Additionally, science majors would take a comprehensive science section including physics, chemistry and biology, while social science majors would take a comprehensive social science section including history, politics and geography. In sum, the social science major requires much less knowledge in math and science.

Moreover, major choice in high school restricts students' choice of major in college[8]. When choosing their majors in high school, students need to balance what they would like to do in the future and what they have a comparative advantage in because of the competitive CEE. Parents and teachers may also play a role when students choose majors.

[8] For instance, a social science student is not allowed to major in physics in college, while a science student is not allowed to major in history. Several majors don't have restrictions, such as traditional Chinese medical science and library management. Policies vary overtime and across universities. Economics major, for example, was restricted to social science student years ago and now accepts applications from both majors.

1.2.2 High School in the Data Set

This research focuses on one public high school in China's Hebei Province. It is located in a city with a population of 1.1 million. Students admitted by this high school are the top 20 percent in the whole province, and over 95 percent of them attend college. The school currently has more than 3,500 students. The data consists of five cohorts beginning with the class of 2011 through the class of 2015. On average, this school has about 1,200 incoming students each year.

Upon entering, students are tracked into different classrooms based on their rankings in the HSEE. According to conversations with the school administrators, the tracking threshold is determined after students have been admitted. On average, 350 students are allocated to the high track. Since the HSEE is not comparable across years, the thresholds vary over time. Moreover, according to Ma and Shi (2014), even though many schools track students, this practice is illegal. Therefore, the threshold is not public information.

Students are arbitrarily assigned into classrooms within each track. They stay in the assigned classroom, and teachers come to give lectures. Students have a fixed group of classmates. After being in a track for one year -- at the beginning of 11^{th} grade -- students choose their majors, change classrooms accordingly, but continue to stay in the same track for another two years.

The workload in high school is substantial. The school day lasts from 7 a.m. to 7 p.m. for six days a week with a two-hour lunch break each day. Students' relative performance in class is readily observable. The ranking of students is common knowledge. Usually, when returning the test sheets, students' names are called from the highest score to the lowest score. Students also receive a paper report for the midterm and final exam with their total score, total rank, scores of five individual fields and rank respectively on them.

1.3 EMPIRICAL APPROACH

1.3.1 Research Design

The challenge of identifying the impact of tracking is that tracking assignment is not random. Tracking schools allocate students into different tracks according to academic performance, which provides a natural setup for a regression discontinuity (RD) design to test whether marginal students are better off being assigned to the high track. This paper adopts a fuzzy RD Design that has been used in literature (Duflo et. al. 2011; Ma and Shi 2014). My empirical approach exploits the fact that HSEE is the main determinant of which track students are assigned to. Marginal students below the threshold should provide counterfactual outcomes for students right above the threshold who were assigned into the high track, since the treatment is effectively randomized in a neighborhood of the threshold.

The distance between initial test scores and the threshold is calculated in the following equation:

$$d_{it} = (total\ \ score_{it} - threshold_t)/s.d._t \tag{1}$$

where $total\ score_{it}$ is the admission score for student *i* at year *t*; $threshold_t$ is the minimal score for high track at year t ; $s.d._t$ is the standard deviation for test score at year *t*.

The reduced form estimation equation is given by:

$$Y_{is} = \beta_0 + \beta_1 * eligibility_i + f(d_{it}) + eligibility_i * f(d_{it}) + \beta_3 X_i + \varepsilon_{it} \qquad (2)$$

where Y_{is} is the academic outcome for student *i* in the end of semester *s*, $eligibility_i$ is the binary indicator of whether student *i*'s score is above threshold or not, d_{it} measures the difference between total score and cut-off point and enters the equation in polynomial form, and X_i is the full set of covariates included in the data set, including gender, year fixed effects and the initial score of each subject. The coefficient of interest is β_1, which measures the impact of scoring above the threshed on test score.

There are non-compliers in the sample: some students in the high track have test score below the cutoff and a few students in the low track have test score above the cutoff. If parents have a strong relationship with the school, then their kids could be assigned to the high track even with test scores lower than the threshold. According to the conversation with the school, it is very unlikely that students with scores above the threshold are in low track. It might be a discrepancy in the data.

Since there are non-compliers in the sample, I also use eligibility as an instrument and estimate the following two stage least square equations:

$$treated_i = \alpha_0 + \alpha_1 * eligibility_i + f(d_{it}) + eligibility_i * f(d_{it}) + \alpha_2 X_i$$

(3)

$$Y_{is} = \gamma_0 + \gamma_1 * treated_i + f(d_{it}) + treated_i * f(d_{it}) + \gamma_2 X_i \qquad (4)$$

When estimating Equation (2) (3) and (4) for boys and girls separately, I employ a linear specification or quadratic specification for the polynomials, $f(d_{it})$. I also utilize local linear non-parametric regression (LLR) with optimal bandwidth around the threshold and triangle kernel (Imbens and Kalyanaraman 2012). The triangle kernel used in the estimation puts more weight on observations closer to the threshold point, which is different from unweighted regressions. According to Imbens and Kalyanaraman (2014), estimators should be based on local linear or quadratic polynomial or other smooth function instead of global high-order polynomials in regression discontinuity analysis[9]. I show that the estimated results are consistent under both the quadratic specification and LLR. I choose the parametric estimation as my preferred specification because it allows me to test the heterogeneous impact by gender.

1.3.2 Data Set

The data set is administrative data obtained from the registrar's office and includes five cohorts from 2008 to 2012. For each student, it contains the score on each subject on the HSEE, finals test scores for the first semester, track and class assignment, major, gender, teacher's assignment and teacher's gender. The scores on the HSEE are missing for students in 2008. Instead, it contains the ranking for the HSEE and I use ranking as the running variable for that year. To test for robustness, I drop 2008 in the appendix tables and reach similar conclusions. Approximately 1,200 students were

---

[9] They point out that global high-order polynomials leads to noisy estimates, sensitivity to the degree of the polynomial, and poor coverage of confidence intervals.

admitted each year[10]. On average, 350 students are allocated into high track with the exception of 2008, in which only 130 students are in the high track. Students from rural areas are dropped from the sample since they are assigned to neither the high track nor the low track[11].

Table 1.1 summarizes the variables used in this analysis by track. Initial score is the standardized HSEE score within each year. Students in the high track have on average 1.3 standard deviation higher scores than students in the low track. There are more female students in the high track. The next two variables, class size and female teachers are calculated at the student level instead of at the class level. Students in the high track have slightly bigger classes and are less likely to have female teachers compared to students in the low track.

The first outcome of interest is the choice of major after one year of tracking. Table 1.1 indicates that the fraction of students who major in science in high track is significantly higher than that in low track. The second outcome of interest is test scores after being tracked for one semester. The final exam contains five sections: three sections each with 150 points: Chinese, math and English; two sections each with 300 points: science and social science, which are created by the school and are the same for both tracks. The tests are graded anonymously by teachers, usually with one section by one

[10] In 2010, it has record of 1200 students in final exams. However, the record for HSEE is incomplete. It only contains the scores of first 800 students ranked by initial test score. Since around 350 students are tracked into high track, the truncated data would provide a reasonable number of students around the threshold, *i.e*. all students in high track, and around 50 percent students in low track. According to the school, although the reason for the missing data is unclear, there is nothing special about this cohort. It has similar number and ability groups as in other years.

[11] Students from rural area are assigned to a different track and live in the school.

teacher. Therefore, the final scores are assumed to be objective measures of individuals' academic achievements. Table 1.1 shows that students in the high track have, on average, 1.28 standard deviation higher final scores than students in the low track. Appendix Table 1.1 summarizes the data for boys and girls separately. Compared to boys, girls have slightly higher initial scores, lower outcome scores and a lower fraction majoring in science.

1.4 INTERNAL VALIDITY

1.4.1 First stage regression

In order to implement the RD design, the assignment to the treatment must vary discontinuously at the cutoff point. I first demonstrate that the treatment is effective in the sense that students are "correctly" allocated into tracks based on the threshold. Figure 1.1 presents the relationship between the probability of being treated (high track) and the distance to cutoff. (a) is for girls and (b) is for boys. $treated_i$ is a binary outcome that equals 1 if individual i is tracked into a high ability class, and equals 0 otherwise. The vertical axis represents the average probability of being assigned to the high track for 40-point bins. The horizontal axis represents the standardized distance from the threshold. The right side of the threshold mainly has probability equal to 1 while the left side predominantly has probability equal to 0. Figure 1.1 suggests there is a jump at threshold for both genders.

Table 1.2 presents the effect of tracking on choice of major. In Table 1.2 Panel A, I estimate the discontinuity in the probability of being allocated in the high track for girls using Equation (3). The discontinuity estimates from column (1) and (2) — the preferred quadratic specification of distance —imply that students have scores above the threshold are about 79 percentage points more likely to be allocated into the high track. Column (3) - (4) estimates a linear specification with narrow bandwidth. Column (5) presents the estimated result using LLR. The first stage estimates for boys are presented in Table 1.2 Panel B. Approximately, 76 percentage points of students are treated based on their eligibility. The analysis above demonstrates that the experiment is valid.

1.4.2 Continuity check

One critical assumption of RD Design is that students are not able to manipulate the treatment and I should not observe a discontinuity in the distribution of the forcing variable (distance to the cutoff). It is a reasonable assumption given that the threshold is determined after HSEE and is not public information. In Figure 1.2, I run a McCrary (2008) test to check for such discontinuity for 2009-2012. As discussed in the data section, students in 2008 only have rankings. Therefore, McCrary test cannot be performed for 2008. Figure 1.2 suggests that the distribution of the forcing variable is continuous around the threshold.

I further test whether initial subjects test scores are continuous around the threshold. Since the total score of school entrance test determines tracking assignment, I expect no individual subject to significantly predict the probability of being assigned to high track. If, for example, despite of similar total scores, marginal students above the threshold have much higher math scores than marginal students below the threshold, then

I suspect that rather than equally weighting subjects, the tracking rule assigns more weight to math. As a result, it may not invalidate the continuity assumptions, but make students less comparable around the threshold and change our understanding of the impact track assignment[12].

Table 1.3 presents evidence of the continuity of initial subject test scores. Panel A presents the sample for girls and Panel B presents the sample for boys. For each subject, the first row is the estimated discontinuity using Equation (2) with year fixed effects and a quadratic function of distance from the cutoff. The results show that scores of each subject are continuous. Since subject test scores may be correlated with each other, I allow the errors of these regressions to be correlated and present the estimated coefficients of using seemingly unrelated regressions. In Appendix Table 1.2, for each subject, I estimate the continuity using LLR model, controlling for year fixed effects. The results in both tables suggest that all subjects are continuous at threshold.

1.5 RESULTS

1.5.1 Effects on the choice of major

I first examine whether tracking has an impact on students' choice of major after the first year in high school. Figure 1.3 depicts the relationship between choice of major and the running variable, which is the standardized distance between initial test scores and tracking threshold. Panel A shows the relationship for girls and panel B is for boys.

[12] I also check the continuity of female students around the threshold for the full sample. I regress female on quadratic specifications of distance and LLR. Both estimated coefficients are not significant (with p-value bigger than 0.23). Therefore, the gender of students is continuous around the tracking threshold.

The y-axis is a binary indicator, which equals 1 if the student chooses science. The red line is a LLR fit and the black line is a parametric fit with quadratic specification of distance. The upward slope of the fitted lines on both sides of the threshold indicates that students with higher test score are more likely to choose science as their major. The jump at the threshold indicates that marginal students in the low track are more likely to choose science than are marginal students in high track. Compared to boys, girls are more likely to be affected by tracking when choosing majors, as indicated by the bigger jump at the threshold.

Table 1.4 presents the effect of tracking on choice of major for female students. Panel A presents the reduced form estimates as in Equation (2). Panel B shows the IV estimates as in Equation (4). The outcome variable is the binary indicator of Science major. In column (1), my preferred specification is estimated after controlling for initial performance in each of the five subjects[13]. Column (1) shows that tracking significantly reduces the estimated probability of choosing a science major by 21 percentage points. In Column (2) – (5), the magnitude of estimates is reduced from 21 to 14 percentage points. These results suggest that tracking reduces the probability of choosing to major in science by at least 14 percentage points for girls.

In Table 1.5, the same outcome is estimated for boys. Column (1) indicates that tracking significantly reduces the estimated probability of choosing to major in science major by 7 percentage points under reduced form estimates. Columns (2) – (4) estimate

[13] All five subjects test scores are not co-linear wit the main running variable-total score, because all test scores are standardized. The total score is calculated by adding the five subjects together and then is standardized.

different specifications. Column (5) presents LLR estimates. Overall, results are consistent and suggest that tracking reduces the probability of choosing science major by around 7 percentage points for boys. Panel C presents the p-values for the null hypothesis that the estimated results are the same for boys and girls. The results indicate that under the preferred specification, tracking has statistically different impacts on male and female students.

1.5.2 Effects on test scores

I next assess the effect of tracking on the total score of the first-semester final exam. Figure 1.4 depicts the relationship between the standardized test score and the running variable. The black line is a parametric quadratic function, and the red line is a LLR fit. For female students in Figure 1.4 (a), the fitted curves seem to be continuous. For male students in Figure 1.3 (b), however, the jump at the threshold indicates that marginal students in the high track are benefited from tracking.

In Table 1.6, I report the results when the sample is restricted to female students. Panel A presents the reduced form estimates as in Equation (2). Panel B shows the IV estimates as in Equation (4). Column (1) and (2) are estimated under quadratic specifications of polynomials. Column (3) and (4) are estimated under linear regression with a narrow bandwidth. Column (5) presents the estimate of LLR. The estimates in column 1-5 are statistically insignificant, suggesting an insignificant effect of tracking on test scores for girls.

Table 1.7 shows the effect of tracking on boys. The estimate in Column (1) suggests that tracking increases marginal male students' test scores by 0.16 standard deviations. This estimate is consistent under different specifications. Panel C presents the p-values for the null hypothesis that the estimated results are the same for boys and girls. The results indicate that under the preferred specification, tracking has significantly different impacts on male and female students.

I then test whether tracking affects students' subject test scores. Table 1.8 presents the estimates of the impact of tracking on test scores by subject for female students, using second order polynomials. For marginal female students in the high track, the results in Table 1.8 suggest that tracking significantly increases Chinese, English and social science test scores. Table 1.9 presents the estimates of the impact of tracking on test scores by subject for boys. For marginal male students in high track, results in Table 1.9 suggest that tracking significantly increases math, science and social science scores.

1.6 MECHANISM

In this section, I explore the mechanism behind the impact of track assignment on choice of major. Compared to similar ability students in the low track, students in the high track have better teachers, higher ability peers and lower relative rank. I also find that their test scores affected by track assignment as well. In sum, I explore the mechanisms in four aspects: test scores, teachers, peers and students' relative ability.

The analysis in this section suggests that being assigned to the high track increases social science but not science scores for marginal attendants. Thus, students in the high track have an advantage in social science and are more likely to choose it as their major. It also indicates that compared to boys, girls in high track are less likely to have same gender teachers in science subjects, and are more likely to be harmed by the competitive study environment.

1.6.1 Test scores

In the Result section, I have shown that although track assignment improves students' total test score, it discourages students from majoring in science. In fact, students may choose majors based on subject test scores instead of their total score. Students in both majors study Chinese, math and English. Science majors have a more demanding math curriculum and study physics, chemistry and biology, while social science majors study geography, history and politics. Since the college entrance exams are different for different majors, students are likely to choose majors based on their test scores.

For example, if track assignment increases test scores on the social science exams relative to test scores on the math and science exam, then students in the high track maybe more likely to major in social science compared to those in the low track. The results in Table 1.8 show that, compared to their counterparts in the low track, girls in the high track increase all subjects except math and science, which leads to an advantage of studying social science. However, the evidence for boys in Table 1.9 is less obvious. Compared to marginal boys in the low track, boys in the high track benefit in math, science and social science, leaving the test score story unclear. Compared with boys, girls are more likely to lose advantage in studying science in high track and they are less likely to major in science.

1.6.2 Teacher effect

In the student-teacher gender match literature, several papers have documented that same-gender teachers might serve as better role models for students (Lim & Meer 2015; Carrington et. al. 2008; Ammermuller & Dolton, 2006; Francis et al., 2006). I expect same gender teacher might be one of the channels since students in high track have higher share of male teachers. In this section, I use the subsample from 2008, 2009 and 2011 that contains teacher's information. It only has teachers' assignment, which subject they teach and the genders. To measure teacher's effect, I use two variables: (1) gender of the teachers, and (2) the share of high track the teacher teaches.

Since science major classes consist of math, physics, chemistry and biology, I expect the teachers for these subject may play a role in explaining the mechanism. In Table 1.10, Column (1)-(4) estimate the difference of female teachers between high and low track using Equation (2) with quadratic specifications. In Panel (A), Column (1) shows that lowest scoring girls in the high track are 15 percentage points less likely to have a female math teacher, compared to the highest scoring girls in the low track. Results for gender of teachers suggest that girls are more likely to have male math and chemistry teachers, and boys are less likely to have female teachers for physics and chemistry. Table 1.10 indicates that compared to girls, boys are more likely to have the same-gender teachers for science subjects. Thus, the gender-match may be an explanation for why girls in high track are less like to major in science compare to boys in the same track.

In order to measure teacher quality, I construct a proxy variable for each teacher using the share of high track classes she or he teaches. In the dataset, most teachers teach more than one class and some of them teach in both high and low track. It is assumed that the higher quality of the teacher, the larger percentage of high track classes she or he teaches[14]. In Table 1.10, Column (5)-(8) shows the results for the quality of teachers. Both boy and girls in the high track have teachers who are more likely to teach high track.

Table 1.11 displays the results for a subsample, which contains teachers' characteristics. Column (1) estimates the main outcome with the same quadratic function and same control variables (year fixed effect and five initial subjects) for the subsample. Column (2) – (5) adds variables for science teachers' gender, quality, and teacher fixed effects. The estimated results for girls in Panel (A) suggest that the magnitude of the impact of tracking decrease slightly after controlling for teacher characteristics. While the results for boys in Panel (B) suggest that the impact of tracking persists after controlling for teacher fixed effect. Overall, this table indicates that difference of teacher characteristics cannot fully explain the difference in majors. There still are unobserved factors that drive the discontinuity in choice of major.

[14] According to the conversion with the school, experienced teachers are more likely to be assigned to high track. In reality, some teachers may select classes as well. Unfortunately, I cannot construct a perfect measure given the limitation of the data.

1.6.3 Peer effect and relative rank

Students in the high track have very different peers and relative rank compared to their counterparts in the low track. Peer group is defined as the students in the same classroom since students spend the entire school day in a fixed classroom. I use two variables to measure characteristics of peers: share of female peers and the mean of initial scores of peers. In addition, I measure the relative rank is the percentile rank for the initial score within each classroom[15]. The higher is the initial test score, the higher is the percentile rank.

Table 1.12 presents the difference in peers and relative rank for students in different tracks. In Panel (A) Column (1), girls in the high track are 13 percentage points more likely to have female peers than girls in the low track. Boys are 10 percentage points more likely to have female peers. The second column suggests that both girls and boys in the high track are more likely to have better quality peers. Panel (A) Column (3) estimates the difference in relative rank for students in high and low track. If all students were tracked by the threshold, then the expected coefficient of *eligibility* is 1. The estimated coefficient is 0.5 because there are some non-compliers.

In order to explore whether these factors affect choice of major, I control for them in the main regression. In Table 1.12, Column (4) estimates Equation (2) under the preferred specification. In the next column, I control for these three variables. The magnitude of the impact reduces by one third for girls while remaining the same for boys. The R-squared for both samples increases to more than 0.8 from 0.18, suggesting that

[15] Relative rank is also defined within each classroom.

these three variables are useful when explaining the choice of major. Peers and relative rank may only play a role when girls choose majors. As discussed in the Introduction, studies (Ho and Kelman 2014; Carrell *et al.* 2010; Dee 2007 and Gneezy *et al*. 2003) have shown that, compared to boys, female students are more likely to be affected by competitive environments. It is plausible that since girls are more likely to be affected by relative rank, overall, they are less likely to choose science majors.

1.7 CONCLUSION

This paper examines the impact of track assignment on students around the tracking threshold. I use the data from a high school in China and exploit an RD design. The outcome variables of interest are students' choice of major in high school and high school test scores. The results suggest that being assigned to high track reduces the probability of choosing science major by 21 and 7 percentage points for marginal female and male students. It also improves test scores by 0.16 standard deviations for boys but its impact on scores is insignificant for girls. The magnitude of the estimated effect is consistent in various specifications.

I show that at the margin tracking exerts heterogeneous impacts on boys and girls. Compared to boys in the high track, girls around the threshold are less likely to major in science and less likely to increase total test scores. In fact, lowest scoring girls in the high track improve their Chinese, English and social science scores while boys increase math, science and social science scores. Since science major is a prerequisite for college STEM majors, the findings suggest that marginal girls in the high track are more likely to be harmed by track assignment in the sense that they are discouraged from majoring in STEM. This finding contributes to the literature of tracking and gender gap in learning.

I explore the possible channels for the impact on major and find that track assignment improves social science for students in the high track. Consequently it might cause fewer students to major in science. The analysis on teacher's effect suggests that, compared to boys, girls have less gender-match teachers in high track. Thus, girls are more affected by the track assignment. I also find that it is plausible that girls are affected by peers and their relative rank when choosing majors.

My findings provide a new dimension of understanding of the impact tracking on choice of major. The causal relationship between tracking and choice of major can apply to other circumstances beyond tracking, such as modeling choice of courses and major in high school and college.

Table 1.1- Summary Statistics

	High track (1)	Low track (2)	Difference (3)
Initial score	0.83	-0.48	1.30***
(stdev.)	(0.50)	(0.92)	(0.02)
Female students	0.57	0.52	0.06***
	(0.50)	(0.50)	(0.01)
Class size	60.46	57.93	2.53***
	(7.09)	(9.00)	(0.24)
Female teachers	0.61	0.65	-0.04***
	(0.11)	(0.14)	(0.01)
Science major	0.75	0.59	0.16***
	(0.44)	(0.49)	(0.01)
Outcome score	0.81	-0.47	1.28***
(stdev.)	(0.65)	(0.88)	(0.02)
Observations	1,799	3,399	

Notes: This table presents the summary statistics for the key variables. Standard deviations are in parentheses. Test scores are standardized. The share of female teachers is calculated using a subsample since teacher's information is missing for 2009 and 2012. In the subsample, there are 653 students in high track and 1,736 students in the low track.
***Significant at 1 percent level
**Significant at 5 percent level
*Significant at 10 percent level

Table 1.2-The Impact of Passing Exam on Admittance to High Track

	(1)	(2)	(3)	(4)	(5)
Panel A: Girls					
eligibility	0.79***	0.78***	0.71***	0.68***	0.70***
	(0.02)	(0.02)	(0.02)	(0.07)	(0.06)
R-square	0.80	0.80	0.73	0.54	---
Mean of Y	0.37	0.37	0.54	0.64	0.37
Observations	2,954	2,954	1,749	409	2,954
Panel B: Boys					
eligibility	0.76***	0.77***	0.65***	0.52***	0.68***
	(0.02)	(0.02)	(0.03)	(0.07)	(0.02)
R-square	0.78	0.78	0.73	0.46	0.78
Mean of Y	0.32	0.32	0.52	0.63	---
Observations	2,542	2,542	1,306	323	2,542
Individual Controls	YES	NO	YES	YES	YES
Running variable	Quadratic	Quadratic	Linear |distance| <0.8	Linear |distance| <0.15	LLR

Notes: This table presents estimates of the first stage results by equation (3). Eligibility is an indicator for students with test scores above the threshold. The dependent variable is a binary indicator, which equals 1 if the student is observed in the high track. Control variables are year fixed effect and initial scores for each five subjects. Panel A is the estimates for girls and Panel B is the results for boys. Standard errors are in parentheses.

***Significant at 1 percent level

**Significant at 5 percent level

*Significant at 10 percent level

Table 1.3-Balance of Observable Characteristics Around Cutoff

	Chinese	Math	English	Science	Social Science
	(1)	(2)	(3)	(4)	(5)
Panel A: Girls					
eligibility	0.03	0.04	0.03	0.03	-0.02
	(0.07)	(0.07)	(0.06)	(0.05)	(0.06)
R-square	0.45	0.61	0.62	0.75	0.65
Observations	2,954	2,954	2,954	2,954	2,954
Panel B: Boys					
eligibility	0.08	0.06	0.06	-0.07	0.06
	(0.09)	(0.08)	(0.08)	(0.06)	(0.07)
R-square	0.50	0.61	0.68	0.76	0.64
Observations	2,542	2,542	2,542	2,542	2,542

Notes: This table presents evidence of the continuity of the individual level characteristics with respect to the distance with year fixed effects and a quadratic specification for the distance function. The estimated standard error of the estimate is in parentheses.
***Significant at 1 percent level
**Significant at 5 percent level
*Significant at 10 percent level

Table 1.4-Impact of Tracking on Choice of Science Majors for Girls

	(1)	(2)	(3)	(4)	(5)
Panel A:					
Reduced form					
eligibility	-0.21***	-0.20***	-0.15***	-0.18**	-0.14***
	(0.03)	(0.04)	(0.04)	(0.09)	(0.05)
R-square	0.24	0.15	0.16	0.20	---
Panel B:					
IV estimates					
treated	-0.28***	-0.27***	-0.22***	-0.27**	-0.23***
	(0.05)	(0.05)	(0.06)	(0.13)	(0.07)
R-square	0.23	0.12	0.17	0.20	---
Mean of Y	0.54	0.54	0.64	0.60	0.54
Observations	2,954	2,954	1,749	409	2,954
Individual Controls	YES	NO	YES	YES	YES
Running variable	Quadratic	Quadratic	Linear |distance| <0.8	Linear |distance| <0.15	LLR

Notes: This table presents estimates of the discontinuity in the relationship between whether female students choose science major and the distance. Control variables are year fixed effect and initial scores for each five subjects. Panel A is the reduced form results given by equation (2) and panel B estimates equation (4). Standard errors are in parentheses.

***Significant at 1 percent level

**Significant at 5 percent level

*Significant at 10 percent level

Table 1.5-Impact of Tracking on Choice of Science Majors for Boys

	(1)	(2)	(3)	(4)	(5)
Panel A:					
Reduced form					
eligibility	-0.07*	-0.08**	-0.07*	-0.12	-0.06**
	(0.04)	(0.04)	(0.03)	(0.07)	(0.03)
R-square	0.23	0.14	0.08	0.11	---
Panel B:					
IV estimates					
treated	-0.08*	-0.09*	-0.08	-0.24*	-0.10**
	(0.05)	(0.05)	(0.05)	(0.15)	(0.05)
R-square	0.22	0.14	0.08	0.08	---
Panel C					
Ho: Boys=Girls					
P-value	0.00	0.00	0.07	0.71	---
Mean of Y	0.75	0.75	0.85	0.86	0.75
Observations	2,542	2,542	1,306	323	2,542
Individual Controls	YES	NO	YES	YES	YES
Running variable	Quadratic	Quadratic	Linear |distance| <0.8	Linear |distance| <0.15	LLR

Notes: This table presents estimates of the discontinuity in the relationship between whether male students choose science major and the distance. Control variables are year fixed effect and initial scores for each five subjects. Panel A is the reduced form results given by equation (2) and panel B estimates equation (4). In Panel C, p-values indicate the confidence level that tracking has same effect on majors for boys and girls can be rejected. Standard errors are in parentheses.
***Significant at 1 percent level
**Significant at 5 percent level
*Significant at 10 percent level

Table 1.6-Impact of Tracking on Test Scores for Girls

	(1)	(2)	(3)	(4)	(5)
Panel A:					
Reduced form					
eligibility	0.04	0.06	-0.02	-0.05	-0.10
	(0.04)	(0.04)	(0.05)	(0.13)	(0.07)
R-square	0.73	0.71	0.54	0.31	---
Panel B:					
IV estimates					
treated	0.04	0.07	-0.03	-0.08	-0.17
	(0.06)	(0.06)	(0.06)	(0.21)	(0.12)
R-square	0.72	0.71	0.54	0.29	---
Mean of Y	-0.01	-0.01	0.38	0.52	-0.01
Observations	2,776	2,776	1,640	247	2,776
Individual Controls	YES	NO	YES	YES	YES
Running variable	Quadratic	Quadratic	Linear |distance| <0.8	Linear |distance| <0.1	LLR

Notes: This table presents estimates of the discontinuity in the relationship between total scores and distance for girls. Control variables are year fixed effect and initial scores for each five subjects. Panel A is the reduced form results given by equation (2) and panel B estimates equation (4). Standard errors are in parentheses.
***Significant at 1 percent level
**Significant at 5 percent level
*Significant at 10 percent level

Table 1.7-Impact of Tracking on Test Scores for Boys

	(1)	(2)	(3)	(4)	(5)
Panel A:					
Reduced form					
eligibility	0.16***	0.17***	0.14***	0.16	0.16***
	(0.06)	(0.06)	(0.06)	(0.14)	(0.06)
R-square	0.70	0.68	0.52	0.30	---
Panel B:					
IV estimates					
treated	0.21***	0.22***	0.20***	0.25	0.28***
	(0.07)	(0.07)	(0.08)	(0.22)	(0.10)
R-square	0.70	0.68	0.52	0.30	---
Panel C:					
Ho:Boys=Girls					
P-value	0.07	0.01	0.02	0.30	---
Mean of Y	-0.08	-0.08	0.46	0.60	-0.08
Observations	2,421	2,421	1,273	208	---
Individual Controls	YES	NO	YES	YES	YES
Running variable	Quadratic	Quadratic	Linear |distance| <0.8	Linear |distance| <0.1	LLR

Notes: This table presents estimates of the discontinuity in the relationship between test scores and the distance for male students. Control variables are year fixed effect and initial scores for each five subjects. Panel A is the reduced form results given by equation (2) and panel B estimates equation (4). In Panel C, p-values indicate the confidence level that tracking has same effect on majors for boys and girls can be rejected. Standard errors are in parentheses.

***Significant at 1 percent level

**Significant at 5 percent level

*Significant at 10 percent level

Table 1.8-Impact of Tracking on Subject Test Scores for Girls

	Chinese	Math	English	Science	Social Science
	(1)	(2)	(3)	(4)	(5)
Panel A: Reduced form					
eligibility	0.15**	0.02	0.13***	-0.06	0.11**
	(0.07)	(0.06)	(0.05)	(0.05)	(0.05)
R-square	0.33	0.51	0.57	0.66	0.59
Panel B: IV estimates					
treated	0.19**	0.02	0.16**	-0.08	0.12*
	(0.08)	(0.07)	(0.06)	(0.06)	(0.07)
Fstats	127.89	266.05	253.18	495.37	404.86
R-square	0.33	0.51	0.57	0.66	0.58
Mean of Y	0.17	-0.08	0.17	-0.08	-0.01
Observations	2,776	2,776	2,776	2,776	2,776

Notes: This table presents the estimates of the discontinuity in the relationship between test scores by subject and distance for girls. The outcomes are five subjects test scores at the end of first academic year. Panel A is the reduced form results given by equation (2) and panel B estimates equation (4). Standard errors are in parentheses.
***Significant at 1 percent level
**Significant at 5 percent level
*Significant at 10 percent level

Table 1.9-Impact of Tracking on Subject Test Scores for Boys

	Chinese	Math	English	Science	Social Science
	(1)	(2)	(3)	(4)	(5)
Panel A: Reduced form					
eligibility	0.05	0.20***	0.09	0.12*	0.18***
	(0.08)	(0.07)	(0.07)	(0.06)	(0.07)
R-square	0.33	0.49	0.52	0.65	0.52
Panel B: IV estimates					
treated	0.06	0.25***	0.11	0.16**	0.23***
	(0.10)	(0.09)	(0.09)	(0.08)	(0.09)
R-square	0.33	0.49	0.52	0.64	0.52
Mean of Y	-0.24	0.02	-0.25	-0.01	-0.09
Observations	2,421	2,421	2,421	2,421	2,421
F-stats	115.85	218.01	217.61	404.78	259.62

Notes: This table presents the estimates of the discontinuity in the relationship between test scores by subject and distance for boys. The outcomes are five subjects test scores at the end of first academic year. Panel A is the reduced form results given by equation (2) and panel B estimates equation (4). Standard errors are in parentheses.
***Significant at 1 percent level
**Significant at 5 percent level
*Significant at 10 percent level

Table 1.10-Difference in Teacher Genders and Qualities by Tracks

	Outcome 1: female teachers				Outcome 2: teachers' quality			
	Math	Physics	Chemistry	Biology	Math	Physics	Chemistry	Biology
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Girls								
eligibility	-0.15***	-0.01	-0.13**	0.03	0.38***	0.19***	0.28***	0.13***
	(0.05)	(0.05)	(0.06)	(0.06)	(0.03)	(0.03)	(0.03)	(0.02)
R-square	0.13	0.18	0.05	0.07	0.45	0.42	0.46	0.50
Mean of Y	0.58	0.47	0.58	0.52	0.38	0.35	0.35	0.35
Observations	1880				1880			
Panel B: Boys								
eligibility	-0.04	-0.21***	-0.19***	0.02	0.30***	0.19***	0.25***	0.14***
	(0.06)	(0.06)	(0.06)	(0.06)	(0.04)	(0.03)	(0.03)	(0.02)
R-square	0.16	0.17	0.06	0.05	0.40	0.38	0.46	0.41
Mean of Y	0.58	0.50	0.63	0.53	0.33	0.34	0.32	0.31
Observations	1665				1665			

Notes: This table presents estimates of the discontinuity in the relationship between teachers' characteristics and the distance. Control variables are year fixed effect and initial scores for each five subjects. Panel A is the reduced form results for girls given by equation (2) and panel B is the results for boys. Standard errors are in parentheses.
***Significant at 1 percent level
**Significant at 5 percent level
*Significant at 10 percent level

Table 1.11-Impact of Tracking on Choice of Science Major

	Science (1)	Science (2)	Science (3)	Science (4)	Science (5)
Panel A:					
Girls					
eligibility	-0.13**	-0.10**	-0.11**	-0.10**	-0.10***
	(0.05)	(0.05)	(0.05)	(0.05)	(0.02)
R-square	0.25	0.27	0.34	0.38	0.94
Mean of Y	0.54				
Observations	1880				
Panel B:					
Boys					
eligibility	-0.04	0.02	-0.02	-0.01	-0.06***
	(0.05)	(0.05)	(0.05)	(0.05)	(0.02)
R-square	0.21	0.26	0.27	0.32	0.94
Mean of Y	0.75				
Observations	1665				
Individual Controls	YES	YES	YES	YES	YES
Female teacher		YES		YES	
Teacher quality			YES	YES	
Teacher FE					YES

Notes: This table presents estimates of the discontinuity in the relationship between choice of science major and the distance for students. Control variables are year fixed effect and initial scores for each five subjects. Panel A is the reduced form results for girls given by equation (2) and panel B is the results for boys. Standard errors are in parentheses.
***Significant at 1 percent level
**Significant at 5 percent level
*Significant at 10 percent level

Table 1.12-Difference in Peers and Relative Rank by Tracks

	Female peers (1)	Peer ability (2)	Relative rank (3)	Science (4)	Science (5)
Panel A: Girls					
eligibility	0.13***	0.96***	-0.50***	-0.21***	-0.14***
	(0.01)	(0.03)	(0.02)	(0.03)	(0.02)
R-square	0.14	0.73	0.54	0.24	0.84
Mean of Y	0.57	-0.04	0.53	0.54	0.54
Observations	2954				
F-stats	238.90	267.58	406.23	495.22	483.56
Panel B: Boys					
eligibility	0.10***	0.94***	-0.56***	-0.06*	-0.08***
	(0.01)	(0.03)	(0.02)	(0.04)	(0.02)
R-square	0.18	0.70	0.60	0.22	0.81
Mean of Y	0.50	-0.02	0.46	0.75	0.75
Observations	2542				
F-stats	214.89	246.53	401.09	469.45	477.28
Controls	YES	YES	YES	YES	YES
Female peers	---	---	---		YES
Peer ability	---	---	---		YES
Relative rank	---	---	---		YES

Notes: This table presents estimates of the discontinuity in the relationship between choice of science major and the distance for students. Control variables are year fixed effect and initial scores for each five subjects. Panel A is the reduced form results for girls given by equation (2) and panel B is the results for boys. Standard errors are in parentheses.
***Significant at 1 percent level
**Significant at 5 percent level
*Significant at 10 percent level

Appendix Table 1.1: Summary Statistics by Gender

	High track (1)	Low track (2)	Difference (3)
Panel A:			
Girls			
Initial score	0.85	-0.41	1.24***
	(0.49)	(0.88)	(0.03)
Class size	59.87	57.26	2.61***
	(7.19)	(9.64)	(0.34)
Female teachers	0.58	0.67	-0.09***
	(0.11)	(0.13)	(0.01)
Science major	0.66	0.48	0.18***
	(0.47)	(0.50)	(0.02)
Outcome score	0.79	-0.47	1.26***
		(0.84)	(0.03)
	(0.63)		
Observations	1,026	1,751	
Panel B:			
Boys			
Initial score	0.81	-0.56	1.32***
	(0.52)	(0.95)	(0.04)
Class size	61.25	58.65	2.60***
	(6.87)	(8.20)	(0.33)
Female teachers	0.61	0.70	-0.09**
	(0.10)	(0.13)	(0.01)
Science major	0.86	0.70	0.15***
	(0.35)	(0.46)	(0.02)
Outcome score	0.84	-0.48	1.32***
	(0.66)	(0.92)	(0.04)
Observations	773	1,648	

Notes: This table presents the summary statistics for the key variables by gender. Standard deviations are in parentheses. Test scores are standardized. The share of female teachers is calculated using a subsample since teacher's information is missing for 2009 and 2012.

***Significant at 1 percent level

**Significant at 5 percent level

*Significant at 10 percent level

Appendix Table 1.2-Continuity Check using LLR

	Chinese (1)	Math (2)	English (3)	Science (4)	Social Science (5)
Panel A: Girls					
eligibility	0.05	0.01	0.01	0.02	-0.01
	(0.06)	(0.06)	(0.04)	(0.04)	(0.05)
Observations	2,776	2,776	2,776	2,776	2,776
Panel B: Boys					
eligibility	0.06	0.04	0.05	-0.02	0.05
	(0.07)	(0.07)	(0.05)	(0.04)	(0.05)
Observations	2,542	2,542	2,542	2,542	2,542

Notes: This table presents evidence of the continuity of the individual level characteristics with respect to the distance with year fixed effects. The coefficients are estimated by the local linear function with triangle kernel estimated on each side of the threshold. The estimated standard error of the estimate is in parentheses.

***Significant at 1 percent level

**Significant at 5 percent level

*Significant at 10 percent level

Appendix Table 1.3-Impact of Tracking on Choice of Science Majors for Girls

	(1)	(2)	(3)	(4)	(5)
Panel A:					
Reduced form					
eligibility	-0.19***	-0.18***	-0.14***	-0.15**	-0.13***
	(0.03)	(0.04)	(0.04)	(0.08)	(0.05)
R-square	0.24	0.14	0.16	0.20	---
Panel B:					
IV estimates					
treated	-0.25***	-0.22***	-0.21***	-0.25**	-0.22***
	(0.05)	(0.05)	(0.06)	(0.13)	(0.07)
R-square	0.23	0.12	0.16	0.20	---
Observations	2,302	2,302	1,547	350	2,302
Individual Controls	YES	NO	YES	YES	YES
Running variable	Quadratic	Quadratic	Linear |distance| <0.8	Linear |distance| <0.15	LLR

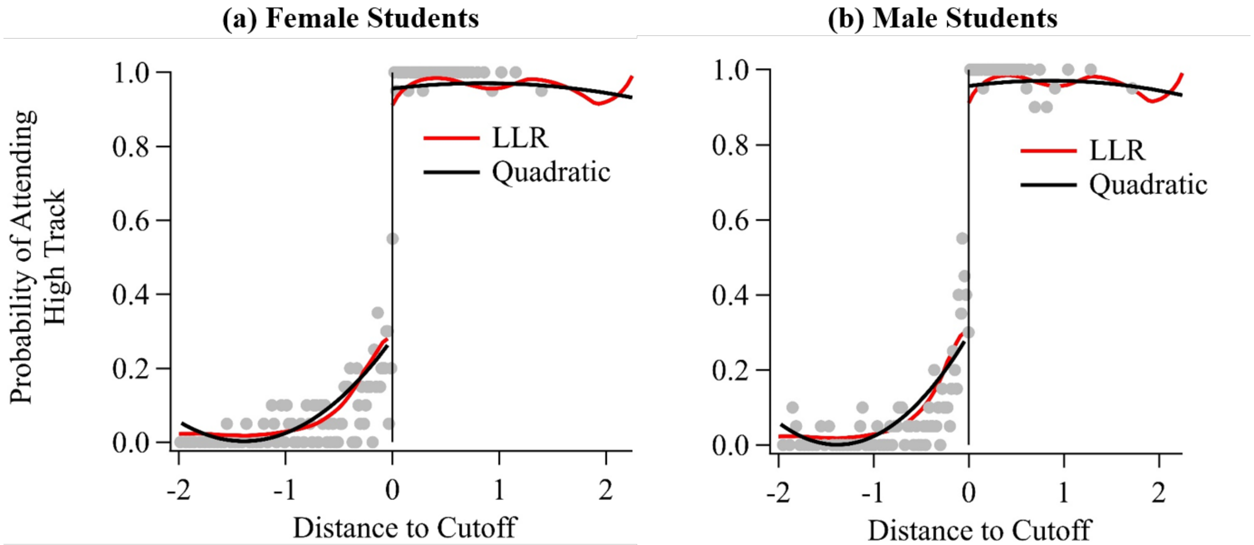
Notes: This table presents estimates of the discontinuity in the relationship between whether female students choose science major and the distance for year 2009-2012. Control variables are year fixed effect and initial scores for each five subjects. Panel A is the reduced form results given by equation (2) and panel B estimates equation (4). Standard errors are in parentheses.
***Significant at 1 percent level
**Significant at 5 percent level
*Significant at 10 percent level

Appendix Table 1.4-Impact of Tracking on Choice of Science Majors for Boys

	(1)	(2)	(3)	(4)	(5)
Panel A:					
Reduced form					
eligibility	-0.08**	-0.08**	-0.07*	-0.12	-0.06***
	(0.04)	(0.04)	(0.03)	(0.07)	(0.03)
R-square	0.22	0.14	0.09	0.11	---
Panel B:					
IV estimates					
treated	-0.09**	-0.09*	-0.09	-0.22	-0.10***
	(0.05)	(0.05)	(0.06)	(0.15)	(0.04)
R-square	0.21	0.14	0.08	0.09	---
Panel C					
Ho: Boys=Girls					
P-value	0.00	0.00	0.06	0.69	---
Observations	1,953	1,953	1,135	279	1,953
Individual Controls	YES	NO	YES	YES	YES
Running variable	Quadratic	Quadratic	Linear |distance| <0.8	Linear |distance| <0.15	LLR

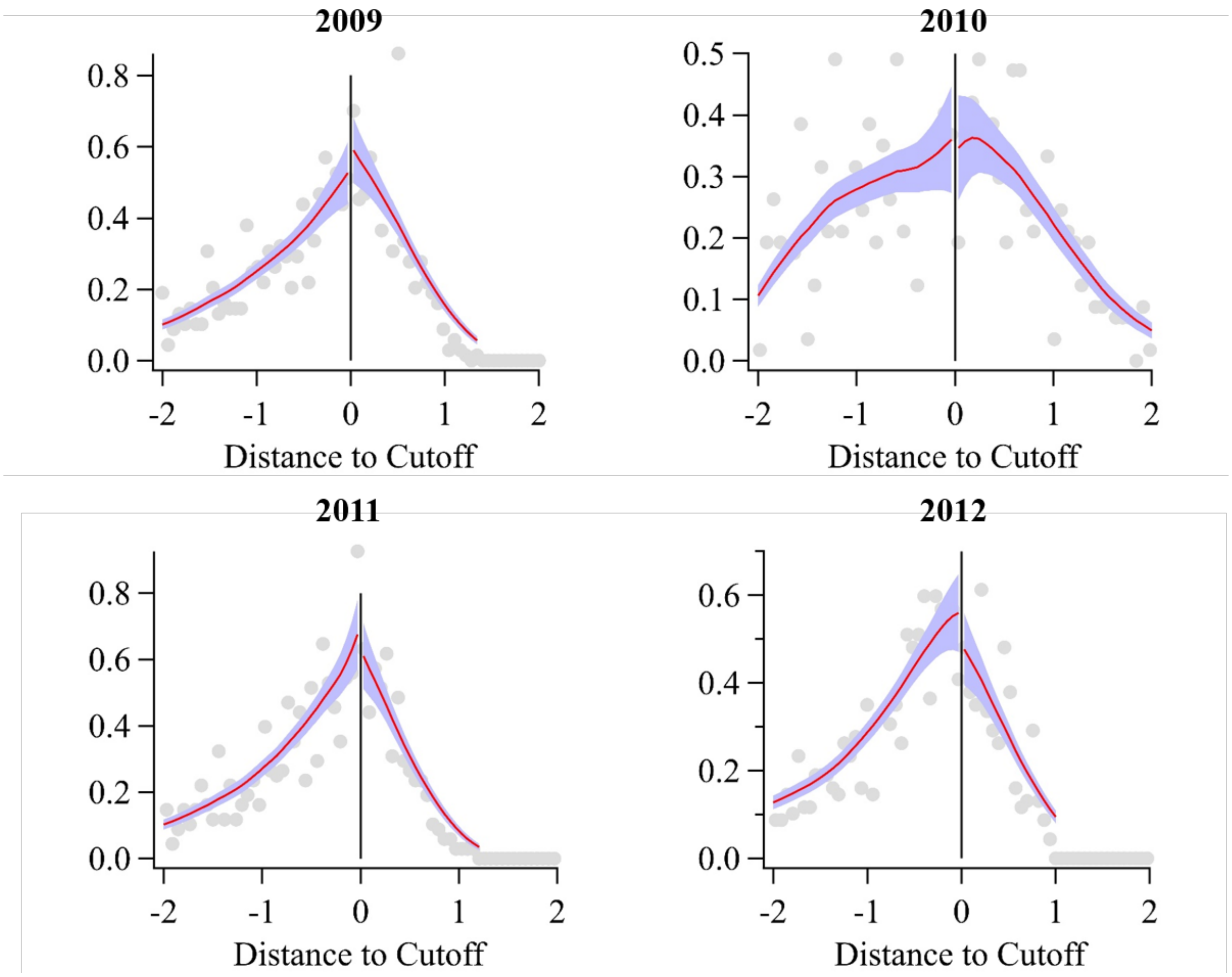
Notes: This table presents estimates of the discontinuity in the relationship between whether male students choose science major and the distance for year 2009-2012. Control variables are year fixed effect and initial scores for each five subjects. Panel A is the reduced form results given by equation (2) and panel B estimates equation (4). In Panel C, p-values indicate the confidence level that tracking has same effect on majors for boys and girls can be rejected. Standard errors are in parentheses.
***Significant at 1 percent level
**Significant at 5 percent level
*Significant at 10 percent level

Figure 1.1- Probability of Being Allocated to High Track.



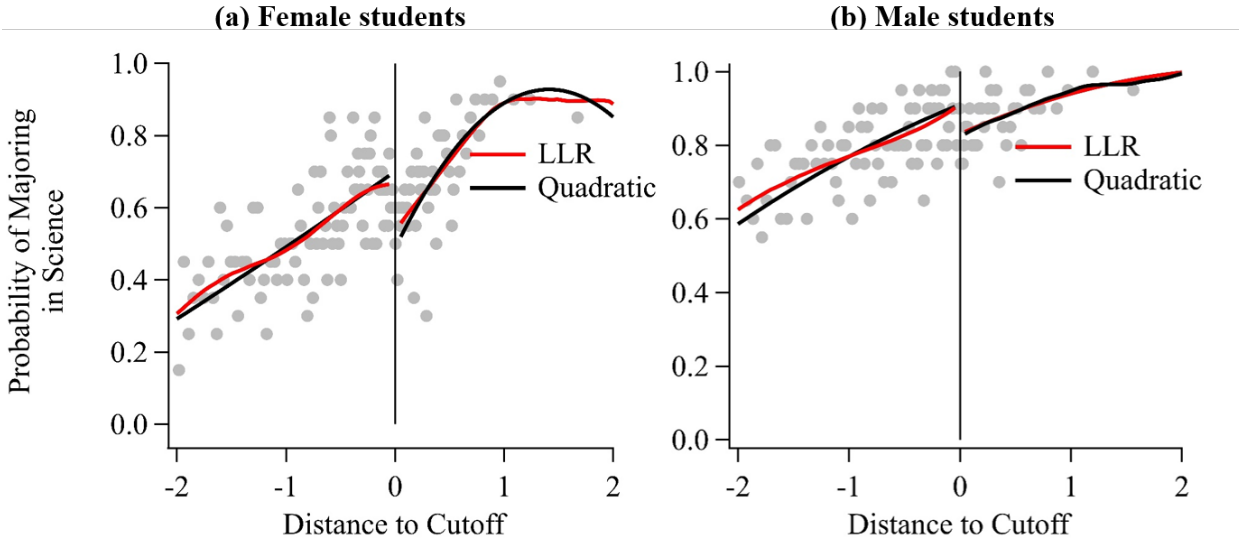
Notes: Figure 1.1 presents the relationship between the probability of being treated (high track) and the distance to cutoff. The x-axis is the standardized difference between entrance exam score and the tracking threshold, and y-axis is the probability of being admitted to the high track. Cutoff for attending high track is represented by the vertical line. Scatter points are the fraction of students attending high track for 40-point bins. Red lines are fitted by the local linear function with triangle kernel estimated on each side of the threshold. Black lines are fitted by quadratic functions of distance.

Figure 1.2-McCrary Test



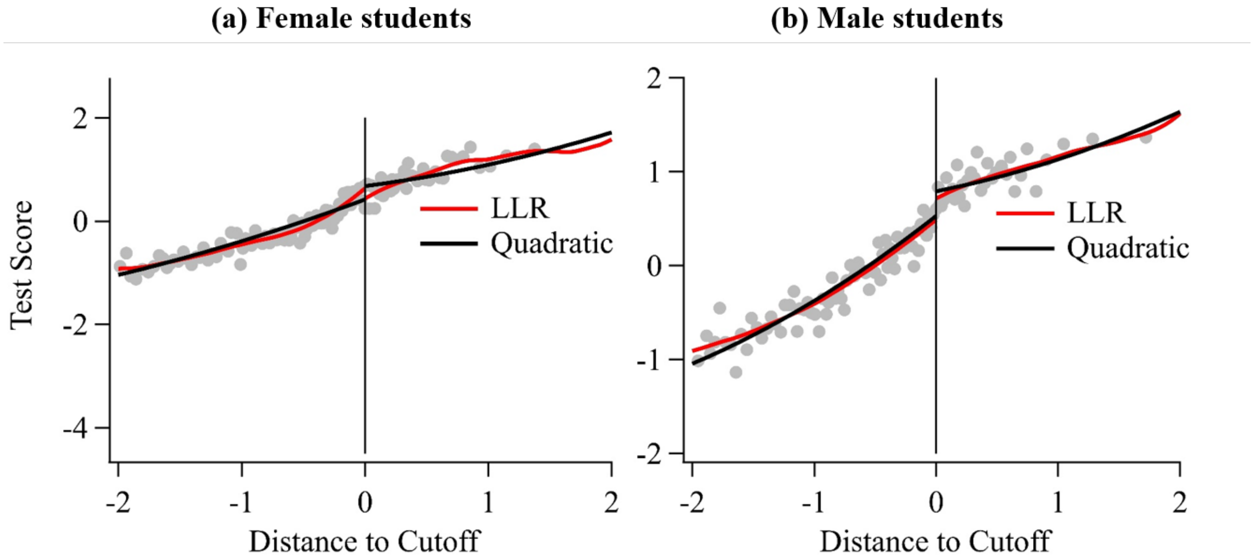
Notes: Figure 1.2 presents the density smoothness test (McCrary 2008) for the standardized distance from admission test score to the cutoff. The red line is the density plot for the distance to cutoff and the band is the 95% confidence interval. It plots the density for four years separately. The distribution in 2010 is different than the distributions of other years because the data only contains the top 66.7% of students' test scores. The bottom one third of observations is missing.

Figure 1.3-Impact of tracking on choosing science major



Notes: Figure 1.3 presents the relationship between the probability of majoring in science and the distance to cutoff. The x-axis is the standardized difference between entrance exam score and the tracking threshold, and y-axis is the probability of choosing the science major. Cutoff for attending high track is represented by the vertical line. Scatter points are the fraction of students with science majors for 40-point bins. Red lines are fitted by the local linear function with triangle kernel estimated on each side of the threshold. Black lines are fitted by quadratic functions of distance.

Figure 1.4-Impact of tracking on test score



Notes: Figure 1.4 presents the relationship between the standardized test score and the distance to cutoff. The x-axis is the standardized difference between entrance exam score and the tracking threshold, and y-axis is the standardized test scores. Cutoff for attending high track is represented by the vertical line. Scatter points are the average score for 40-point bins. Red lines are fitted by the local linear function with triangle kernel estimated on each side of the threshold. Black lines are fitted by quadratic functions of distance.

Chapter 2. Does international student enrollment affect whether domestic students major in STEM fields?

2.1 Introduction

Increasing the number of students in Science, Technology, Engineering, and Math (STEM) majors has become, for many business and policy leaders, a key strategy for keeping the U.S. competitive in a global economy and for keeping the nation's pace of innovation (Carnevale, Smith and Melton, 2011). Currently, 40% of bachelor's degree earned by men and 29% earned by women are in the STEM majors[16]. A 2012 report by the President's Council of Advisors on Science and Technology suggested that in order to meet the growing demand for STEM workers in the US work force, the number of STEM graduates needed to increase by 34 percent over the next decade.[17]

In fact, many students intend to major in a STEM field when they enter college. However, the persistence rate for STEM majors is very low, especially for women and minorities (Griffith 2010). Understanding how and why students decide whether or not to persist in a STEM major is critical to crafting policies that effectively increase the number of STEM graduates.

[16] Report: Growth in Science and Engineering (S&E) Bachelor's Degrees by Gender

[17] PCAST, Engage to Excel: Producing One Million Additional College Graduates with Degrees in Science, Technology, Engineering, and Mathematics (PCAST, Washington, DC, 2012).

The total number of international students in the US higher education system has doubled over the last decade. (See Figure 2.1, which shows the numbers of international students studying in the U.S. over the past fifty years.) Nearly one million international students studied in the U.S. during the 2014-2015 academic year; more than 41.6% of them majored in a STEM field[18]. International students play an important role in the U.S., not only because they contribute more than 30.5 billion dollars and more than 373,000 jobs to the U.S. economy,[19] but also because, as college students, teaching assistants, and instructors, they may affect the learning environment and educational choices of domestic students.

The goal of this chapter is to investigate whether the presence of international student discourage or encourage domestic students to select a STEM major. To answer this question, I use the Texas Educational Research Center (ERC) administrative data from 1994 to 2013, which covers detailed enrollment and graduation data for those who attended college in Texas. The ERC data also provides information about international students' country of origin.

There are several ways international students might either positively or negatively affect domestic students' educational outcomes. International students bring above-average revenue (funding per student) to the school, so that the school may have more resources to benefit all students. Several papers have documented that increased international student enrollment leads to more, or at least not diminished, domestic

[18] Open doors: Fast Facts 2016
[19] Source: NAFSA International Student Economic Value Tool

student enrollment (Zhou 2011, Machin and Murphy 2014). However, if the short-run supply of education resources is inelastic, international students can also directly compete with and crowd out domestic students. Reduced education resources per student due to rising cohort size are therefore likely to negatively affect the number of domestic students at a school, and consequently the supply of college-educated domestic workers entering the STEM labor market (Bound and Turner 2007).

Peer effects can also impact domestic students' decisions to study STEM fields. International students are very likely to major in STEM and tend to earn the top grades in college mathematics (Barnett et al. 2004) [20]. On one hand, high-achieving international students may increase both the quality of domestic students' education and their interest in studying STEM. Ost (2010) found that low-ability students benefited from exposure to stronger peers in science classes. On the other hand, domestic and international students may compete intensely in STEM classes for good grades, and domestic students with lower grades may be less likely to pursue a major in STEM fields. Luppino and Sander (2012) found that lower-ability, non-minority students typically respond to greater competition in the sciences by shifting their choices of major. Several other studies found that some students may leave STEM field majors because of an unpleasant experience in the college Calculus I class (Rasmussen and Ellis 2013; Crisp, Nora, and Taggart, 2009). Fischer (2017) found that women who enrolled in a class with higher-ability peers are less likely to graduate with a STEM degree, while men's STEM persistence is unaffected.

[20] Barnett et al. (2004) find that foreign students did better than the non-immigrant students in by 5.9 points on average out of 100 points.

A major obstacle to identifying the causal relationship between international student enrollment and domestic students majoring in STEM fields is the endogeneity of foreign enrollment. Specifically, if there are some school-level unobservable factors that attract simultaneously more international students and more domestic students who major in STEM fields, the share of international students will be positively correlated with the number of domestic STEM degrees. An ordinary least square approach would therefore estimate only correlation, not causality. To solve this endogeneity problem, I use an instrumental variable, the historical share of international students, to predict the current share of international students. This identification strategy has been widely used in immigration-related literature and papers that examine college enrollment patterns (Card 2005; Machin and Murphy, 2014; Orrenius and Zavodny, 2015).

When examining the impact of international students on domestic students, it is important to consider not only the total number of international students but also the composition of that student body. Both the number and the average ability of international students may vary over years. For example, selective universities may keep their number of admitted international students fixed but recruit higher-ability students over time as the total number of international applicants increases. To better understand peer effects, I explore two different ways of measuring the ability of international students. First, I group international students by the official language (English vs. non-English) of their countries of origin. Students from non-English speaking countries might have comparative advantages in STEM fields compared to disciplines that require English language skills. Meanwhile, if domestic students in STEM field majors are exposed to an increasing number of international students who speak poor English, they may find it more rewarding to switch to a major that rewards English skills. Second, I will use a comparable measure of math and science scores of students from each foreign country to calculate the average scores for international peers in each school. I expect the ability of international peers, especially in math and science, to affect domestic students' decisions regarding whether or not to major in STEM.

This paper finds that a one percentage point increase in the share of international students from non-English speaking countries increases the likelihood of majoring in STEM by 0.9% for domestic male students, while it decreases the probability of majoring in STEM by 1.3% for domestic female students. The impact also differs by domestic ethnic group: it is negative for minorities and insignificant for whites. The results are similar when I group international students by their test scores.

To date, only Orrenius and Zavodny (2015) have examined the effect of international student enrollment on domestic student enrollment to STEM majors. Their findings suggested that when the share of international students in the student body increases by 10%, the probability of domestic female students majoring in science or engineering fields drops by 0.25 percentage points, while the effect is insignificant for domestic male students. My estimated effects show similar trends but a much larger amplitude.

This research has three advantages compared to the Orrenius and Zavodny (2015) study. First, I measure the actual share of international students in each college, while the previous paper only approximates the share by using the population of 18- to 22-year-old immigrants living in the same region as a given college. Second, the dataset I use provides much more information, including a set of pre-college characteristics with high school standardized test scores and ACT scores. I can therefore control for these characteristics and college fixed effects, making my estimates more precise. Third, I examine the impact of international students not only by their total numbers but also by their English language, math, and science abilities, which enables me to better explore the mechanisms by which international student enrollment affects domestic students' decisions about STEM majors.

The rest of this chapter is structured as follows: the next section presents the estimation strategy. Section 2.3 describes the dataset and variables used in the analysis. Section 2.4 presents the estimation results. Section 2.5 concludes the chapter.

2.2 EMPIRICAL STRATEGY

I use a linear probability model to examine the relationship between whether domestic college graduates majored in STEM and two measures of international students: one is the share of international students from a non-English speaking country; the other is the average ability of international students, measured by the Program for International Student Assessment (PISA) average scores for math and sciences for each international student's country of origin.[21] The basic regression model for the first measure is:

$$Y_{ist} = \alpha_1 Share_{st} + X_{ist} + \tau_t + \varphi_s + error \qquad (1)$$

[21] For more detail about these measures, see data section.

The dependent variable Y_{ist} is a binary indicator which equals one if individual *i* at school *s* in cohort *t* graduated from college with a STEM major. $Share_{st}$ is the share of international students from non-English speaking countries for school *s* and cohort *t*. X_{ist} is a set of individual level covariates including gender, race (black, Hispanic, Asian, and other mutually exclusive categories, with non-Hispanic whites as the omitted category) and ACT scores. These variables control for systematic differences in the probability of majoring in STEM across genders, races, and abilities.

The regression also includes school and cohort fixed effects. The school fixed effects φ_s controls for unobservable factors that are specific to a school but constant over time, such as selectivity, research facilities, and enrollment. The cohort fixed effect τ_t controls for unobservable factors that are specific to a college cohort, such as labor market conditions and the change of popularity of STEM majors. The standard errors are robust and clustered on school.

If $Share_{st}$ is not correlated with the error term, we can interpret α_1 as the impact of international students on the number of domestic STEM graduates. However, $Share_{st}$ is endogenous since factors that affect the share of international students may also directly affect the educational choice of domestic students. For example, increasing investment in STEM resources within a university over time may simultaneously attract more international students and more domestic students who want to study STEM. In this case, the ordinary least squares (OLS) estimates will have an upward bias.

I exploit an instrumental variable (IV) to control for the potential endogeneity of the share of international students. The IV used in this project is the historical share of international students in each campus, which has been used in previous literature (Card 2005; Machin and Murphy, 2014; Orrenius and Zavodny, 2015). Specifically, the IV is:

$$Predicted_{st} = \frac{\frac{\sum_c N_{cst_0}}{\sum_s \sum_c N_{cst_0}} * \sum_s \sum_c N_{cst}}{Total_{st}} \tag{2}$$

And the first stage regression is:

$$Share_{st} = \beta_1 Predicted_{st} + X_{ist} + \tau_t + \varphi_s \tag{3}$$

where N_{cst} is the number of foreign students at school s time t from non-English speaking country c. $\sum_s \sum_c N_{cst}$ is the total inflow of international students from non-English speaking countries at each year. $\sum_c N_{cst_0}$ is the total number of international students at school s in base year t_0 from all non-English speaking countries. $Total_{st}$ is the total enrollment at school s in time t.

The instrument is valid only if it affects international student enrollment, but does not directly affect domestic students' choice of major. The historical pattern is a good IV under the network theory assumption: International students have the tendency to attend universities that enrolled a large number of students from the same country in the past because the ethnic networks that already existed in those universities can reduce the informational and mental costs associated with foreign study. For example, if one school historically enrolled a large share of Japanese students, then current applicants from Japan may find this school more attractive than other Texas universities, because it is easier for them to get information about this university and they can expect more help from the Japanese student network upon enrollment.

The network theory suggests that a university with better-established networks by international students from certain countries will see a larger increase in demand for its education when the total number of students from those countries increases compared with a university with fewer connections to those countries. As long as the patterns of foreign enrollment do not change overtime and country-specific networks affect international students' choice of universities, this variable will be positively correlated with the observed foreign enrollment.

This instrument also requires that historical inflows of international students are not systematically related to unobservable factors that affect whether or not current U.S. natives major in STEM. This is likely to be true given that there are several years between the base time and the samples in this analysis. The instrument I created represents the average historical enrollment pattern of international students in universities from 1994 to 1997, while the samples in this study are from 2002 to 2009.

As mentioned in the introduction, both the number and the average ability of international students may affect the outcomes of domestic students. Therefore, I also examine the impact of international students' ability on probability of graduating with STEM degrees for domestic student:

$$Y_{ist} = \alpha_1 Ability_{st} + X_{ist} + U_{st} + \tau_t + \varphi_s + error \qquad (4)$$

where $Ability_{st}$ is the average PISA score for international students in campus *s* at time *t*. U_{st} is the total number of international students from countries that do not participate in PISA testing. Both $Ability_{st}$ and U_{st} are instrumented similarly by the following IVs:

$$Predicted\_Ability_{st} = \frac{\frac{\sum_{c'} N_{c'st_0} * T_{c'}}{\sum_s \sum_{c'} N_{c'st_0} * T_{c'}} * (\sum_s \sum_{c'} N_{c'st} * T_{c'})}{total\ \ enrollment_{st}} \qquad (5)$$

$$Predicted\_U_{st} = \frac{\frac{U_{st_0}}{\sum_s U_{st_0}} \sum_s U_{st}}{Total_{st}} \qquad (6)$$

And the first stages are:

$$Ability_{st} = \beta_1 Predicted\_Ability_{st} + X_{ist} + \tau_t + \varphi_s \qquad (7)$$

$$U_{st} = \gamma_1 Predicted\_U_{st} + X_{ist} + \tau_t + \varphi_s \qquad (8)$$

where $T_{c'}$ is the average score of science and math for country c'. The only difference in notation between these two measures is that $N_{c'st_0}$ is the total number of students from any country c' (including both English speaking countries and non-English speaking countries).

2.3 DATA

The primary individual level data used in this work comes from the Texas Education Research Center (ERC). ERC data covers cohorts in the academic years between 1994 and 2013. I also use country-level cognitive test scores from PISA and measures of cognitive skills constructed by Hanushek and Woessmann (2012).

The college application data includes all students who applied to public universities in Texas and covers cohorts from 2000 and onward. It includes students' ACT or SAT score and admission status. College data covers the time period from 1994 to 2013 for public schools and 2002 to 2013 for private universities. It contains students' enrollment status, college majors, type of degree, type of tuition, and many other individual characteristics, including country of origin. Table 2.1 lists the available variables for different time periods. Due to the limited information on private schools in the data, I will focus on students in public schools for the rest of the chapter.

International students sampled by the ERC came from more than two hundred countries. Figure 2.2 presents the trends of both total enrollment and international student enrollment for public schools from 2002-2013. The enrollments are normalized relative to the first year values. Total enrollment increases over the entire ten-year period. International student enrollment, however, first slightly decreased and then increased, and reached the same level in 2013 as in 2002.

When examining the impact of international students, I use the share of students from non-English speaking countries instead of the total share of international students for several reasons. First, as discussed in the introduction, students from non-English speaking countries are more likely to have a comparative advantage in STEM fields. Therefore, I expect the estimated effect to be larger and more precise. Second, the variation of the total number of international students is very small across time for public schools (Figure 2.2, the major part of the dataset for the analysis). If I use total number of international students in Equation (1), then the variation of the IV only comes from the change of inflow, which is almost constant over time.

I first define English-speaking countries, according to Bleakley and Chin (2004), as those from which more than half of the recent adult immigrants did not speak a language other than English at home. Bleakley and Chin (2004) investigated immigrants from 30 English-speaking countries and 56 non-English speaking countries, which are a subset of the countries in my dataset. For the rest of countries in my dataset, I define

English-speaking countries as those where English is the official language according to the World Factbook. Figure 2.3 depicts the change in the share of students from non-English speaking countries in public universities from 2012-2013. In 2002, 68% of international students came from non-English speaking countries. This number rose to 80% by 2013.

I then construct the measure of international students' cognitive skills based on the work of Hanusek and Woessman (2012). They proposed a consistent measure for math and science skills for countries from 1972-2003. Since the sample in this paper covers the period 1993-2009, I rescale PISA 2006 and 2009 test scores for math and science according to the Hanusek and Woessman algorithms and thereby extend the period of their measure to 2009. The modified measure of cognitive skills contains 76 countries and covers 75% of the international students in the data set. Figure 2.4 presents the share of students from the countries without PISA scores. It does not fluctuate much over time and I instrument for it in the regression as well. Figure 2.5 presents the average math and science skills for international students over time, which decreased from 4.59 (out of 5) in 2002 to 4.53 in 2009.

The regression sample only includes degree-seeking students enrolled in college from the years 2002 to 2009. Students who enrolled before 2002 are dropped from analysis. Since I use data from 1994 to 1997 to construct the base year prediction, I need to keep at least four years[22] between the base year and the years that the IV predicts to

[22] The students used to construct the IV should not be the same students who are in the regression sample. I assume that students would take four years to graduate.

make the IV valid. Since this paper focuses on the effect of international students on the number of domestic students with STEM degrees, the regression sample is restricted to students who actually graduated. Table 2.2 presents the summary statistics for the share of international students and graduates in public schools. Only 3% of the full sample were international students. This number increases to 5% if the sample is restricted to students with STEM majors. The proportion of international students in the graduation data is similar. Table 2.3 presents the summary statistics for all variables used in the regressions. 19% of students graduated with a STEM major, 8% of whom were engineering majors, 10% of whom were science majors, and 1% of whom were math majors.

2.4 RESULTS

This section reports the estimated results of the relationship between graduating in a STEM field in college and two international student measures: share of international students from non-English speaking countries and average math and science skills for international students. Both measures approximate the ability of international peers, which may affect the likelihood of obtaining a STEM degree for domestic students.

2.4.1 Non-English speaking country

Table 2.4 reports the first stage results from Equation (3). The IV is constructed by Equation (2). The table shows five specifications and each coefficient comes from a separate regression: the first column is for the entire sample; the next two columns contain subsample by gender; and the last two columns are subsamples by race. For each specification, I control for test score, school, and time fixed effects. The first column indicates that if historically the number of students from non-English speaking countries increases by 1, then the predicted number of these students increases by 1.57. The results in Table 2.4 show that the correlation between the IV and the actual share of international students is bigger than zero (F-test statistics above 20), suggesting that the IV is valid.

Table 2.5 presents the two stage least square results estimated from Equation (1). Each coefficient comes from a separate regression. The top panel shows estimates using the full sample. The first column estimates the impact of international students on the number of domestic students with STEM degrees. The next three columns examine this effect using more specific major categories (science, engineering, and math). In the first panel, I control for gender, race, ACT scores, school, and time fixed effects. In the second and third panels, I control for the same set of covariates except for gender. In the last two panels, I exclude race.

In Table 2.5, the second column in Panel A indicates that a 1% increase in the number of international students from non-English speaking countries reduces the probability of graduating with a science major by 0.584% for domestic students. The effect for math majors is also negative but smaller (-0.118%). In the last column, the effect for engineering majors is positive. Specifically, a 1% increase in the share of international students from non-English speaking countries increases the likelihood of graduating with an engineering major by 0.43% for domestic students. Panel B presents the estimated results for male domestic students and Panel C is for female students. Overall, the effect on two genders is quite different. In particular, with a 1% increase of international peers, female students are 1.3% less likely to major in STEM, while male students are 0.909 % more likely to graduate with a STEM major.

The last two panels from Table 2.5 present the results estimated using either only whites or minorities. The minority group includes both black and Hispanic students. The results suggest that minorities are negatively affected by the increasing share of international peers while the whites are either not affected or do not benefit.

2.4.2 Math and Science skills

Table 2.6 reports the first stage results from Equation (7). It includes five specifications: the first column is for the entire sample; the next two columns contain subsample by gender; and the last two columns are subsamples by race. For each specification, I control for test scores, school, and time fixed effect. The results in Table 2.6 suggest the instrument is valid since its correlation with the actual share of international students is greater than zero (F-test statistics is above 20).

Table 2.7 presents the instrumental variable results estimated from Equation (4). Each coefficient comes from a separate regression. Panel A reports the estimates using the full sample. The first column estimates the impact of international students on the number of domestic students with STEM degrees. The next three columns examine this effect using more specific major categories (science, engineering, and math). I control for gender, race, ACT scores, school, and time fixed effects. In Table 2.7, the second column in Panel A indicates that a 1% increase of average ability of international students significantly reduces the probability of graduating with a science major by 0.137% for domestic students. The peer effect for math majors is also negative (-0.052%), but smaller. The last column shows that, similar to results obtained using the share of non-English speaking peers, the effect for engineering majors is positive and significant. Panel B presents the estimated results for male students and Panel C is for female students. Overall, the effects on men are insignificant while the impacts on women are negative and significant. When there is a 1% increase in the ability of international peers, female students are 0.324% less likely to major in a STEM field.

The last two panels from Table 2.7 present the results estimated using only white and only minority students respectively. The results suggest that minorities are 0.182% less likely to graduate with a science major if their peers' average ability increases by 1%. Whites are less likely to be affected by their better international peers, the exception being the math degree outcomes.

2.4.3 Discussion

The estimates that come from both measures of international peers have a similar direction of effect. That is, students majoring in science and math degrees are likely to be negatively affected either by an increase in students from non-English speaking countries or by an increase in students from countries with better math and science test scores. The impacts are heterogeneous by gender and race: female and minority students experience more negative peer effects, whereas male students either benefit or are not largely affected by their international peers.

There are several mechanisms that may explain the heterogeneous effects by major. For example, compared to engineering majors, science and math majors require less skill in English and communication. Engineering majors, for instance, are more likely to have group projects. Therefore, if international peers have a disadvantage in English language skills in engineering, domestic students may be more likely to switch to majors that reward English skills.

2.5 CONCLUSION

This study examines whether international students affect US domestic students' decision to major in STEM in undergraduate study. I construct an IV based on the historical enrollment pattern of international students to instrument the current enrollment pattern of international students. My findings suggested that one percentage increase in the share of international students from non-English speaking countries increases the likelihood of domestic students majoring in STEM by 0.9% for males and decreasing the probability of majoring in STEM by 1.3% for females. The impact also differs by ethnic group: it is negative for minorities and insignificant for whites. The results are similar when I measure international students by their test scores. Therefore, policies aimed at increasing the number of domestic students with STEM degrees should pay particular attention to female and minority students. In addition, the results may also inform immigration policy.

This research contributes to the existing literature by providing two new measures for the quality of international peers. The first is the share of students from non-English speaking country and the second is the average math and science skills of international peers. These two measures are more relevant to the decision of STEM majors compared to the measure in current literature, which is the total number of international students.

This research has several limitations. First, the ability measure for international students is an average score for an entire country. It is likely that students who study in the U.S. have scores above average in their home country. Moreover, the distribution of tests scores may vary across countries. The measure may therefore underestimate the math and science skills of international students. Second, the data only allows us to define “peers” as all the international students in a given school. In fact, it would be more accurate to define “peers” as the international students in a given class.

Table 2.1-Availability of variables for public and private schools

Variable	Public schools	Private schools
Enrollment and graduation data	1994-2013	2002-2013
ACT score in application data	2000-2013	N/A
Indicator for international students	1994-2013	2002-2013
Country of origin in enrollment data	1994-2013	2007-2013
Country of origin in graduation data	N/A	N/A
Base year data (IV construction)	1994-1997	2003
Data in analysis sample	2003-2009	2006-2009

Table 2.2-Share of international students in the sample

	Enrollment: full sample		Enrollment: subsample	
	Percentage	**Female**	**Percentage**	**Female**
Domestic	97	0.55	95	0.35
International	3	0.44	5	0.29
Non-English speaking	2	0.45	3.4	0.27
	Graduation: full sample		Graduation: subsample	
	Percentage	**Female**	**Percentage**	**Female**
Domestic	97	0.57	94	0.35
International	3	0.43	6	0.26

Notes: Data used in this table covers public schools between 2002-2013. The subsample includes only students with a STEM major/degree

Table 2.3- Summary Statistics

Variables	Mean	S.D.
STEM major	0.19	0.39
Engineering	0.08	0.28
Science	0.10	0.30
Math	0.01	0.11
Female	0.552	0.50
ACT	23.80	4.83
White	0.57	0.50
Black	0.10	0.30
Hispanic	0.23	0.43
Asian	0.07	0.26
Native	0.01	0.07
Other	0.02	0.18

Notes: This table shows the sample mean and standard deviation for students who graduated in at least 6 years from public universities between 2002 to 2009. (N=997,046)

Table 2.4-First stage regression for international students from non-English speaking countries

	Total	**Male**	**Female**	**White**	**Minorities**
IV	1.57**	1.84***	1.41**	1.98**	1.43**
	(0.57)	(0.59)	(0.57)	(0.82)	(0.56)
F-stats	35.912	31.817	33.088	33.156	34.224
N	997,046	447,175	549,871	567,789	343,532

Notes: Each estimated coefficient comes from a separate first stage regression. The dependent variable is the actual share of international students from non-English speaking countries. Each column corresponds to a subsample. Regressions also include controls for gender, race, ACT score, school, and time fixed effects. Standard errors (in parentheses) are robust and clustered on the school level.
***Significant at the 1 percent level
**Significant at the 5 percent level
*Significant at the 10 percent level

Table 2.5-IV Regression estimates of the relationship between number of domestic students with STEM degree and the share of international students

	STEM	**Science**	**Math**	**Engineering**
Panel A: Total				
Share	-0.270	-0.584***	-0.118**	0.430**
	(0.209)	(0.181)	(0.495)	(0.211)
Mean of Y	19.2	9.6	1.1	8.5
Panel B: Male				
Share	0.909**	0.327*	-0.066	0.661*
	(0.426)	(0.179)	(0.080)	(0.340)
Mean of Y	28	11.1	1.4	15.7
Panel C: Female				
Share	-1.310***	-1.311***	-0.159**	0.149
	(0.272)	(0.335)	(0.065)	(0.117)
Mean of Y	12	8.4	1	2.6
Panel D: White				
Share	0.175	-0.196	-0.101*	0.489*
	(0.315)	(0.217)	(0.061)	(0.297)
Mean of Y	18.5	8.6	1.2	8.9
Panel E: Minority				
Share	-0.660**	-0.733***	-0.166***	0.229
	(0.263)	(0.267)	(0.046)	(0.168)
Mean of Y	16.8	8.8	1.1	6.9

Notes: Each estimated coefficient comes from a separate second stage regression. The dependent variable is a binary indicator and corresponds to the column name. Each panel shows the estimated results for a subsample. Regressions also include controls for gender, race, ACT score, school, and time fixed effects. The mean of Y is the percentage of students who graduated with the major in the column name. Standard errors (in parentheses) are robust and clustered on the school level.
***Significant at the 1 percent level
**Significant at the 5 percent level
*Significant at the 10 percent level

Table 2.6-First stage regression for international students' mean math and science skills

	Total	**Male**	**Female**	**White**	**Minorities**
IV	1.316***	1.467***	1.225***	1.682***	1.181***
	(0.329)	(0.363)	(0.321)	(0.589)	(0.242)
F-stats	27. 810	27.092	23.808	23.546	24.215
N	997,046	447,175	549,871	567,789	343,532

Notes: Each estimated coefficient comes from a separate first stage regression. The dependent variable is the mean score of math and science for international students. Each column corresponds to a subsample. Regressions also include controls for gender, race, ACT score, school, and time fixed effects. Standard errors (in parentheses) are robust and clustered on the school level.
***Significant at the 1 percent level
**Significant at the 5 percent level
*Significant at the 10 percent level

Table 2.7-IV regression estimates of relationship between number of domestic students with STEM degrees and average ability of international students

	STEM	**Science**	**Math**	**Engineering**
Panel A: Total				
Share	-0.140	-0.137**	-0.052***	0.044**
	(0.113)	(0.058)	(0.018)	(0.110)
Mean of Y	19.2	9.6	1.2	8.5
Panel B: Male				
Share	0.130	0.032	-0.032	0.123
	(0.190)	(0.055)	(0.025)	(0.151)
Mean of Y	28	11.1	1.4	15.7
Panel C: Female				
Share	-0.324***	-0.241***	-0.066*	-0.020
	(0.080)	(0.108)	(0.017)	(0.085)
Mean of Y	12	8.4	1	2.6
Panel D: White				
Share	0.127	0.054	-0.094*	0.171
	(0.132)	(0.163)	(0.056)	(0.186)
Mean of Y	18.5	8.6	1.2	8.9
Panel E: Minority				
Share	-0.173	-0.182***	-0.044***	0.049
	(0.118)	(0.061)	(0.013)	(0.127)
Mean of Y	16.8	8.8	1.1	6.9

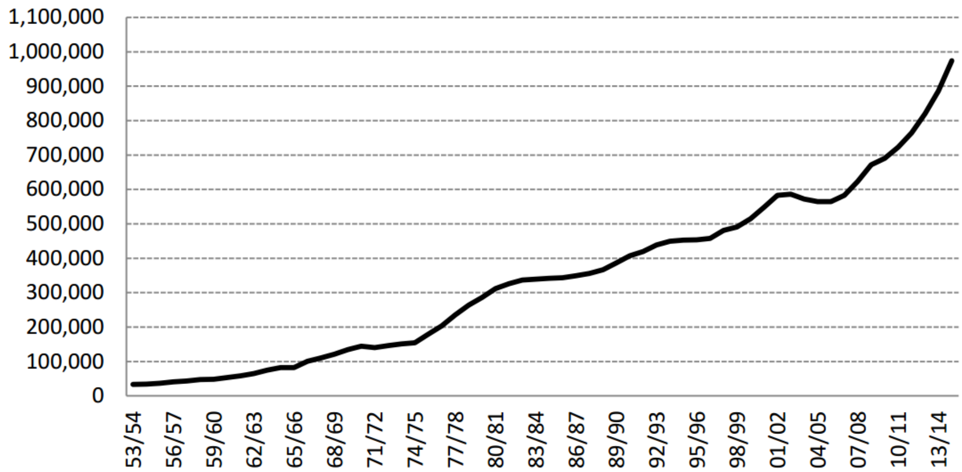
Notes: Each estimated coefficient comes from a separate second stage regression. The dependent variable is a binary indicator and corresponds to the column name. Each panel shows the estimated results for a subsample. Regressions also include controls for gender, race, ACT score, school, and time fixed effects. The mean of Y is the percentage of students who graduated with the major in the column name. Standard errors (in parentheses) are robust and clustered on school level.

***Significant at the 1 percent level

**Significant at the 5 percent level

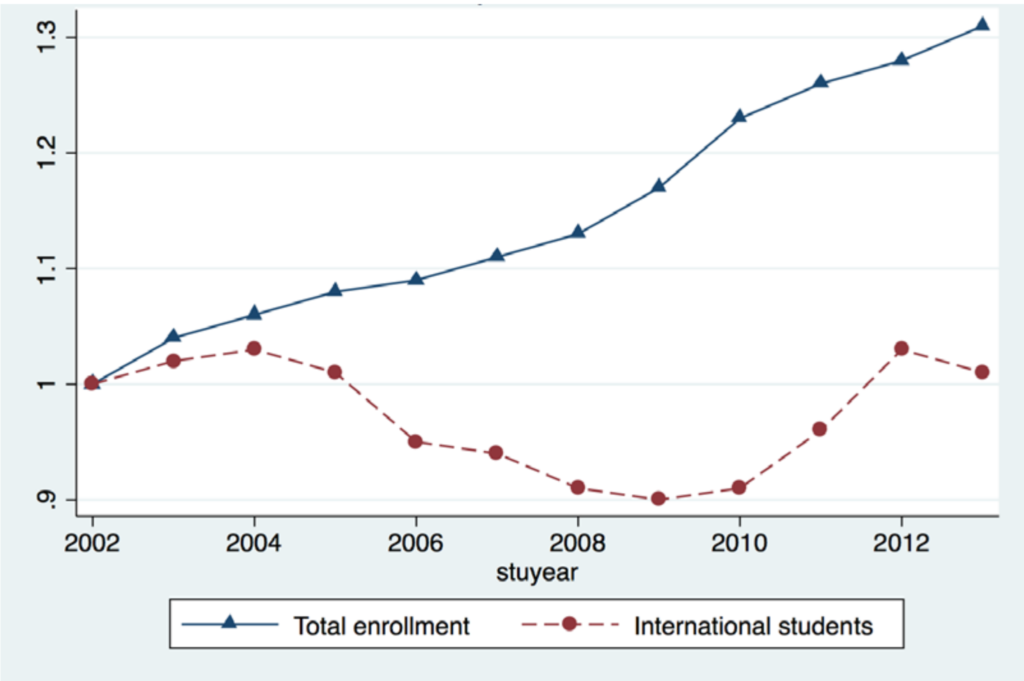
*Significant at the 10 percent level

Figure 2.1- Number of International Students in the US from 1953 to 2013



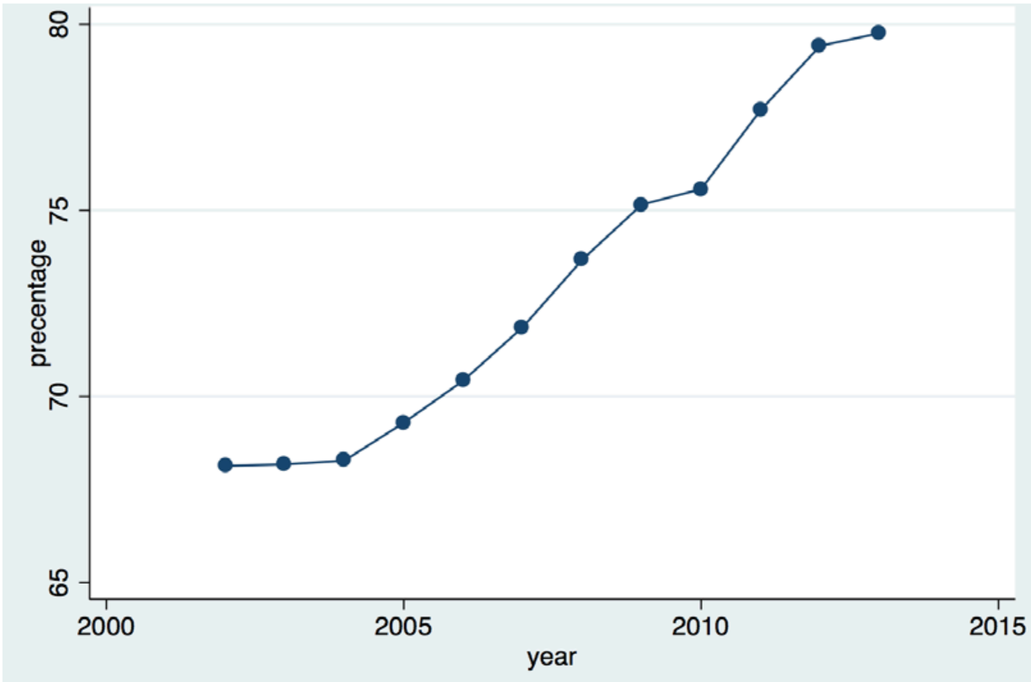
Notes: Figure 2.1 plots the increasing total number of international students in the U.S. postsecondary education system over the past fifty years. Source: Open Door Fast Fact 2016.

Figure 2.2-The College Enrollment Pattern for Texas Public Schools from 2002 to 2013



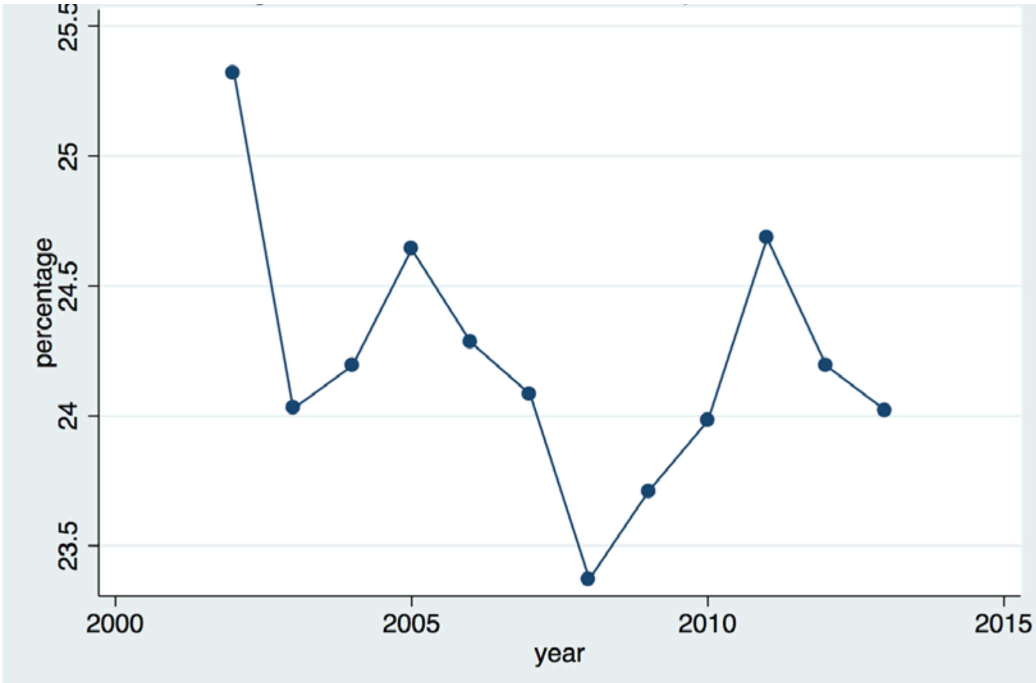
Notes: Figure 2.2 presents total enrollment and total number of international students in Texas public universities from 2002 to 2013. The y-axis is normalized to one relative to the year 2002.

Figure 2.3- The Share of International Students from Non-English Speaking Countries in Texas Public Universities from 2002 to 2013



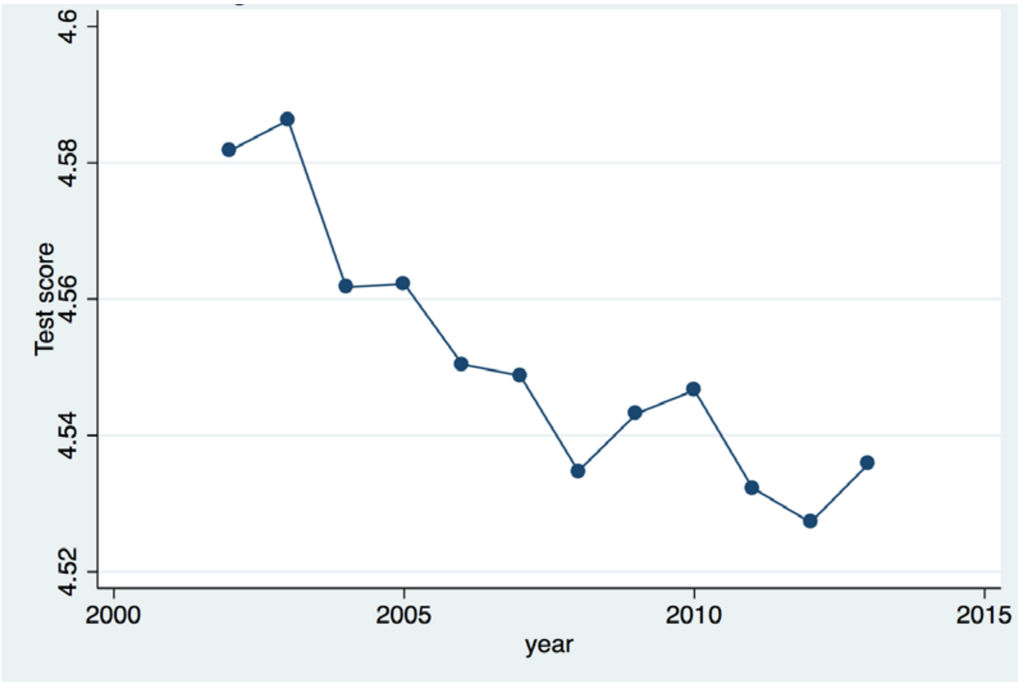
Notes: Figure 2.3 depicts the change in the share of international students from non-English speaking countries at Texas public universities from 2002 to 2013. The x-axis is time and the y-axis is the percentage of international students from non-English speaking countries in the total number of international students.

Figure 2.4-Percentage of Students from Countries Without Test Scores



Notes: Figure 2.4 presents the percentage of international students from countries without PISA scores public universities from 2002 to 2013.

Figure 2.5-The Average Math and Science Test Scores of International Students



Notes: Figure 2.5 presents the average math and science skills for international students from 2002 to 2013.

Chapter 3: The impact of education on religion in China[23]

3.1 Introduction

Education and religion have both been found to have important impacts on individual outcomes. In terms of education, researchers have shown that more years of schooling lead to higher wages (Card 1999; Fang *et al.* 2012), better health conditions (Silles 2009; Powdthavee 2010), and other positive outcomes (Oreopoulos and Salvanes 2009). Similarly, religious participation has been found to reduce the likelihood that individuals engage in risky or harmful behaviors (Hungerman 2010; Osafo *et al.* 2013). There is a belief held among many social scientists that there is a link between educational levels and religious participation (McCleary and Barro, 2006; Hungerman 2014; Meyersson 2009; Brown and Taylor 2007; Gruber 2005; Glaeser and Sacerdote 2001; Cesur and Mocan 2013; Gulesci and Meyersson 2012). However, there is little empirical evidence to support or contradict these predictions.

There are numerous mechanisms that can possibly be responsible for either the positive or negative effects of education on religiosity. On one hand, it may be that religion is derived from irrational human fears and anxieties (Hume 1757), and therefore, it is ignorant for individuals to be religious. This is the view of the secularization hypothesis (an important theory of religiosity), which predicts that higher levels of education lead to lower levels of religious participation and belief. This is because education imparts scientific knowledge, which can undermine the credibility of religions that rely on belief in supernatural forces (McCleary and Barro, 2006). The educational

[23] This paper is coauthored with Jia Xu

climate in China may be particularly adept at weakening religious belief, as students are taught to only believe in Atheism (specifically communism). Given the national gains in education and the educational climate in China, we might find a negative relationship between education and religiosity.

On the other hand, there are several arguments that stand in contrast to the secularization hypothesis. One is that religious beliefs require abstract thinking. Highly educated people are more capable of speculative reasoning that is needed for intellectual inquiry. Therefore, a more educated person may be more religious (Hungerman 2010). Another argument, which is proposed by Sacerdot and Glaseser (2008), is that education increases the returns from networking. More educated people are thus more likely to participate in social activities, including religious activities and community events. Due to these internal and external motivations, education may have a positive effect on religious belief and participation.

Reflecting these conflicting hypotheses, evidence of the direction of the relationship between education and religion is mixed. Researchers have documented both positive (Meyersson 2009; Brown and Taylor 2007; Gruber 2005; Glaeser and Sacerdote 2001) and negative (Hungerman 2014; Cesur and Mocan 2013; Gulesci and Meyersson 2012) correlations between education and religious affiliation. For example, religious trends in North America are consistent with the secularization hypothesis. Both the United States and Canada have witnessed a decrease in religion affiliation from the 1970s into the 2000s, but an increase in average years of schooling (Hungerman 2014). However, unlike North America, China saw increases in both the number of people who identify themselves as religious and years of schooling within the same period. Due to these contrasting societal trends, it is of interest to examine whether the negative effects of education on religion are also present in China, or if there are societal differences in the relationship between education and religiosity (Hungerman 2014).

Additionally, the current literature on this topic has many limitations. Most of the studies that have examined the impact of education on religiosity only use ordinary least squares analysis, which can be vulnerable to omitted variable bias. This is problematic because it is likely that unobserved factors are correlated with both educational attainment and religiosity (Glaeser and Sacerdote 2008). In addition, the literature is focused primarily on developed countries, and relevant studies for developing countries are rare. This is important because many developing countries have cultural and religious backgrounds that are fundamentally different from those in developed countries. For instance, only a small portion of China's population holds religious beliefs (less than 22 percent) as compared to developed countries (for example, around 60 percent in Canada—Hungerman 2014). Conceptions of what constitutes religious belief may also differ, as religions in China have broad definitions and are often mixed with local culture (Du 2010). For this reason, any research that attempts to examine religion must be sensitive to cultural climate, otherwise they may overlook certain types of religious belief.

The primary goal of this paper is to examine the causal relationship between years of schooling and religious beliefs in China. To do so, we use cross-sectional data from the 2007 Spiritual Life Study of Chinese Residents. We exploit the variation in the compulsory school law’s implementation in China, which increased compulsory schooling from six to nine years. The treatment group is the students who completed junior high school under the compulsory school law, and the control group is the students who were not affected by this law. Compared to their peers in the control group, students in the treatment group had more exposure to both politics and science curriculums, which could affect religious belief. Compulsory school laws have been widely used as instrumental variables when studying the causal impacts of education on outcomes such as income and health (Oreopoulos and Salvanes 2009; Fang et al. 2012)

Our study also adds to the existing literature in two ways. First, to our best knowledge, this is the first paper to examine the relationship between education and religious belief in a large, developing country. Second, the definition of religion used in our analysis encompasses nontraditional religious beliefs that provide a more accurate representation of the beliefs and spiritual life of Chinese people.

The main outcome of our interest is whether someone believes in any formal religion. We find that one additional year of education reduces the probability of being religious by 8 percentage points. Given that only 21.5 percent of people believe in formal religion in China, this effect is sizable. Since a significant number of Chinese people who believe in supernaturality but don’t claim to be religious, we also use another measure, which equals one if the individual has any supernatural beliefs. Overall, there are 38.6 percent of people who believe in supernaturality. My findings suggest that one additional year of education reduces the probability of believing in supernaturality by 10.2 percentage points.

The rest of this paper is structured as follows: the next section introduces the current situation of religious beliefs in China. Section 3.3 presents our strategies to estimate the causal impact of education on religion. Section 3.4 describes the data and variables used in the analysis. Section 3.5 presents the estimation results. Section 3.6 describes the placebo test of the instrumental variable, and explores possible mechanisms to explain the results. Section 3.7 concludes this study.

3.2 BACKGROUND

3.2.1 Religion in China

China has seen a steady increase in religious observance over the past forty years that has sometimes been called a "religion revival" by researchers (Du 2010). According to Albert (2015), less than 300 out of 850 million people[24] (35.3%) were religious in China in 1970, but this proportion increased to 800 million out of 1.37 billion people[25][26] (58%) by 2015. Researchers have attributed this increase to the de-regulation of religions after the Cultural Revolution (1966-1976). Although the government party is officially atheist, it has grown more tolerant of religious activities in recent years (Du 2010).

Figure 3.1 depicts the trend of population who believe in at least one of the following five major religions in China: Chinese folk-religion, Buddhism, Christianity,

---

[24] National Bureau of Statistics of China (http://www.stats.gov.cn/ztjc/ztfx/qzxzgcl60zn/200909/t20090911_68637.html). There is no publicly available data about the percentage of religious population over time.

[25] National Bureau of Statistics of China (http://www.stats.gov.cn/tjsj/zxfb/201604/t20160420_1346151.html)

[26] The total population with religious belief is much higher in our paper compared to the number in other documents since we include the local belief while most researcher only include formal religion beliefs. Excluding local belief, the number of religious people increases from around 80 million to 360 million in the past forty-five years.

Islam, and Taoism. Chinese folk-religion (or local beliefs) not only includes elements of Confucianism, Taoism, and Buddhism, but also involves traditions such as belief in spirits and ancestor worship (Cohen 1992). As shown in the figure, the number of people with local beliefs decreased between 1950 and 1970, but then increased after 1970. From this figure, we can also see that there were increases in belief in Buddhism and Christianity. The number of Buddhists increased after 1970 as Buddhism became more popular among educated people (Du 2010). It also appears that Christianity became more popular, likely following the growth of China's economy and its connection to the western world. In all, this figure suggests that there was a general increase in religious belief in China after 1970.

3.2.2 Primary and junior high school education in China

The nine-year compulsory education includes primary school (5 or 6 years) and junior high school (typically 3 years). The school year for primary and junior high school is divided into two semesters: the fall semester that begins in September and ends before Chinese New Year, and the spring semester that begins after Chinese New Year and ends by July (Lam 2011).

The curriculum for primary education [27] includes nine compulsory courses (mathematics, science, social studies, political education, Chinese, physical education, music, fine arts, and labor skills), with a foreign language as an elective. The exit exams for primary school only include Chinese and mathematics. In junior high school, the

[27] The courses and exit tests may vary across provinces.

curriculum includes five courses (mathematics, science, social science, Chinese, and English) that would be tested for high school entrance exam. Social science courses are comprised of history, political science and geography, while Science courses include physics, chemistry, and biology (Lam 2011).

Because high school entrance exams include subjects in science and social science, which are not included in the primary school exit exams, science and social science courses are taken more seriously in junior high school than in the primary school. We believe that the greater emphasis on these courses will likely affect the religious beliefs of students. This paper investigates the students who would have otherwise dropped out of junior high school if there were no compulsory school law. The treatment group is the students who completed junior high school under the compulsory school law and the control group is the students who dropout either from junior high school or primary school before the implementation of this law. Compared to peers in the control group, students in the treatment group have more exposure to both politics and science curriculum.

3.3 ESTIMATION STRATEGY

As we discussed in section 3.2, education and religiosity are both endogenous. An ordinary least square (OLS) model relating education and religion will overestimate the causal effect if omitted variables impact religion and education in the same direction. For example, if one is ambitious and likes social networking, he or she may invest in education and participate in religious activities at the same time. For this reason, it is necessary to alleviate or overcome the endogeneity issue.

Instrumental variables are used in this study to address the endogeneity issue. Specifically, we use the change in China's compulsory schooling law (implemented in July 1986) as our instrument. This law required that all Chinese children receive nine years of free, compulsory schooling from the ages of 6 to 15 (Xi and Mo 2014). The law was adopted at different times depending on the level of economic development in different regions. Cities and economically developed coastal areas implemented it faster than villages and economically underdeveloped areas (Rawlings 2014). We assume that the rollout is not correlated with religion in that region. We use the actual effective date of implementation in each province to measure the impact of the law on years of education. In this paper, we use the compulsory school law as an IV and estimate a two stage least square model:

$$Y_{it} = \beta_0 + \beta_1 Educ_{it} + X_{it} + \gamma_i + \varepsilon_{it} \quad (1)$$

$$Educ_{it} = \alpha_0 + \alpha_1 Treated_{it} + X_{it} + \gamma_i + \epsilon_{it} \quad (2)$$

where Y_{it} is the binary outcome variable that equals 1 if individual *i* at the survey time t had any religious affiliation; $Educ_{it}$ measures the years of schooling; $Treated_{it}$ is a binary indicator, which equals one if the individual was affected by the compulsory school law and equals zero otherwise; X_{it} is a set of covariates including gender, age, minority status, marital status, household registration at the age of 15 and religious affiliations of parents; γ_i is a province fixed effect and ϵ_{it} is the error term. The coefficient of interest is β_1. If education positively affects one's religion-related behaviors, we expect β_1 to be positive.

3.4 DATA AND VARIABLES

We use data from the 2007 Spiritual Life Study of Chinese Residents (SLSC), a national multi-stage probability-proportional-to-size sample of 7,021 individuals. The sample includes 56 locales throughout China, including 3 municipal cities (Beijing, Shanghai, and Chongqing), 6 provincial capitals (Guangzhou, Nanjing, Wuhan, Hefei, Xi'an and Chengdu), 11 regional level cities, 16 small towns, and 20 administrative villages. The final data set is weighted to reflect population parameters in the 2006 Statistical Yearbook of China.

In choosing our control group, we first exclude individuals who were born before 1949 (the year People's Republic China was founded) because of the change of policy environment. According to Fang *et al.* (2012), cohorts who were born in the 1950s experienced disruption in education due to the political turmoil of the Cultural Revolution (1966-1976). To avoid downward bias in schooling within the control group, we further reduce our samples to only those born in 1961 or later, because they were younger than 15 by the end of the Cultural Revolution in 1976. In total, our control group contains 2,019 individuals.

In choosing our treatment group, we restricted our sample to individuals who were aged 12 or younger at the years of 1986. According to the policy, students should be enrolled in school beginning at age 6 and should complete their compulsory education at the age of 15. However, due to regional variation in the implementation of the compulsory education law, individuals who were aged 13-15 years at the time of implementation were not necessarily affected by the law. After we restrict our sample in the manner, 2,394 individuals are included in our treatment sample. Therefore, in all, our sample includes 4,413 individuals.

To measure religiosity, we create a dummy variable equal to 1 if an individual has any religious affiliation, and 0 otherwise. We also study the impact of education on different religions by creating two different binary variables. These two binary variables are *Buddhism*, which is equal to 1 if an individual is a follower of Buddhism and 0 if he/she is not, and *western religion*, which indicates whether individuals report that they believe in at least one of the three religions: Protestantism, Catholicism, or Islam.

The key explanatory variable in our analysis is the number of years of schooling that an individual has completed. We also controlled for other explanatory variables, such as age, gender, ethnic minority, religion of parents, marital status, urban residency at the age of 15, and province. Table 3.1 shows the summary statistics for the sample used in the analysis. The average educational attainment of individuals in the treated group is 1.4 years higher than that of the control group. The treated group is 15 years younger on average than the control group.

Figure 3.2 depicts the distribution of education over different religious affiliation. In the first row, among people without any religious belief, 6.35% have less than a primary school education; 17.06% have only completed primary school; 38.08% have a junior high school education; 27.48% have a high school education; and only 11% have a college education and above.

In addition to organized religion, we also examined the prevalence of spirituality among individuals in our sample, which we consider to be a broader definition of religiosity in this paper. We use the following parameters to create an index of spirituality: whether one thinks religion is important, whether one believes that some western or local gods exist, and a set of variables that measure religion related activities such as praying or wearing a religious item[28]. We will discuss how we define the index to measure spirituality in the Result section.

3.5 RESULTS

This section provides estimates of the relationship between education and religion. Figure 3.3 shows the residual educational attainment for our pre-policy (control) and post-policy (treatment) cohorts. The jump at birth year equals 0 indicates that the policy changes the average years of schooling.

Table 3.2 reports results from the first stage regression. The dependent variable is years of schooling. Results in Table 3.2 suggest that compulsory education law did significantly raise the educational attainment. The first column indicates that students who were affected by the compulsory school law received, on average, 0.835 more years of schooling compared to those in the control group. Columns 2-3 estimate the effect of the law on education levels for males and females respectively, and indicate that the law had a larger effect on females. This educational inequality between the genders may arise from the son preference tradition in China: parents are more likely to invest in their son's education.. The implementation of the compulsory schooling law may have allowed girls to remain in school longer than they would have under normal circumstances. Columns

[28] The list of items are in Q3 and Q4 of footnote 30

4-5 estimate the impact of the law for urban and rural subgroups and show that urban students are more affected.

Table 3.3 presents the estimates of education on religion belief. In the data, individuals were asked if they believe in any of the following religions: Buddhism, Daoism, Confucianism, Catholicism, Islam, and Protestantism. The dependent variable is coded as 1 if one believes in any of these religions. Model 1 is estimated by Equations (1) and (2). Results in the first column suggest that OLS estimate of the relationship between education and religion is not significant. The second column shows the first stage estimate for people who have non-missing values for the dependent variable (that is, for whom we have education data) and religion. The third column shows the second stage estimate. The result suggests that one additional year of schooling significantly reduces the probability of an individual being religious by 8 percent. In Model 2, we add province specific time trends as apart of the control variables and find that the results are similar to those found using Model 1.

In Table 3.4, we estimate the impact of education on five religious measures using Equation (1) and (2). *Buddhism* is defined to be 1 if one believes in Buddhism. *Religious Importance* indicates whether people think religion is important, *Western Supernatural Belief* indicates whether people have one of the following supernatural beliefs: God, heaven, hell, or Jesus Christ. *Local Supernatural Belief* indicates whether people believe in the existence of one of the following supernatural beings: Sages, Ghosts, fate and fortune, god of wealth, and ancestral spirits. Results in Table 3.4 suggest that education has insignificant impact on Buddhism and whether people think religion is important or not. However, education reduces the probability of having western beliefs, western supernatural beliefs, and local supernatural beliefs by 4, 12, and 10 percent, respectively.

We further exploit principal component analysis (PCA) to construct two religious measures: religiosity (which measures whether people have religious objects or behaviors) and attitudes towards religion. Since there are multiple questions in the survey for both measures, the PCA technique is used to construct a single index for each measure and eliminate information redundancy. PCA is a method used to reduce many variables into a few principal components, which can reflect the information provided by the original variables by a linear combination of the original variables. It is widely used to solve the multicollinearity problem for independent variables in regression (Tipping and Bishop 1999). In this paper, we follow the steps of PCA to construct a dependent variable *Religiosity* in regression.

Religiosity is relatively difficult to measure in China compared to western countries since the local belief is not an organized, unified system of beliefs and practices. Chinese people regard western beliefs in a similar way. Chinese, for example, may wear a cross without formally practicing Christianity.[29] Therefore, we complement that measure by adding behavioral evidence available in the dataset.[30]

[29] Similarly, people wear local belief objects may not believe in any gods in local belief. These items may be given by relatives who believe in local belief.

[30] Q1"Have you done the following things? 1. Fortune telling, including face reading; 2. Feng Shui; 3. Analyze one's writing to predict one's destiny; 4. Interpret one's dreams to predict one's destiny; 5. Astrology; 6. Witchcraft; 7. Ask for assistance from someone with pa 8. Automatic writing; 9. Believe in the Eight Diagrams; 10. None of the above".

Q2"Do you believe the existence of soul?".

Q3"Do you have any of the following items at home? 1. Statue or portrait of the god of wealth; 2. Christian objects such as the cross; 3. Muslim objects; 4. Buddhist objects such as statue or portrait; 5. Daoist objects

Panel A of Table 3.5 shows the eigenvalues of the correlation matrix that correspond to each of the principal components. We only keep the first principal components in our analysis, which explains 37% of the variability of the dataset. We then calculate the rotated factor matrix and compute the principal component scores using the eigenvectors. The principal component regression analysis was performed on the shear strength and the scores of the first principal components. Below is the representation of the first principal component:

$$Index_1 = 0.37f_{10}$$

Where $Index_1$ is the shear strength of SFRC beams, andf_{10} is the score of the first principal components. The estimated results are presented in the first column of Table 3.6. The coefficient of *Years of Schooling* is negative, but not significant. Based on this finding, years of schooling don't influence the religiosity of individuals in our sample.

Second, we use PCA to construct an index encoding the attitude about religion based on three questions.[31] As previously mentioned, the compulsory policy in this study

---

such as statue or portrait; 6. Statue or portrait of Confucius; 7. Statue or portrait of Chairman Mao; 8. Ancestral tablets; 9. Statue or portrait of other gods or spirits; 10. No religious objects."

Q4"Do you have any of the following items in your workplace, such as the restaurant or shopping mall? 1. Statue or portrait of the god of wealth; 2. Christian objects such as the cross; 3. Muslim objects; 4. Buddhist objects such as statue or portrait; 5. Daoist objects such as statue or portrait; 6. Statue or portrait of Confucius; 7. Statue or portrait of Chairman Mao; 8. Ancestral tablets; 9. Statue or portrait of other gods or spirits; 10. Inapplicable, no workplace; 11. No religious objects."

[31] Q1"Do you think religion has any negative impact on society?1. Make people irrational; 2. Make people easy to be deceived by fraud; 3. Make people easy to be deceived by fraud; 4. Create conflicts in families; 5. Create social conflicts; 6. Give bad people a chance to defraud others; 7. Corrupt moral standards and social ethology; 8. Waste financial resources; 9. Religion has no negative impact on society"

Q2"People should not have any religious belief; instead people should only believe communism. 1. Agree; 2. Disagree."

aims to increase the years of education from six to nine. During the three years of junior high school, students are taught that communism is the only correct belief to hold. Additionally, it is in this phase of schooling that students start to learn physics, chemistry, and biology beginning in 7th grade. Thus, we may expect the compulsory school law can affect the religious outcomes of students by influencing their attitudes towards communism and science.

Panel B of Table 3.5 shows the eigenvalues of the correlation matrix that correspond to each of the principal components. We keep the first principal components in our analysis, which can explain 51% of the variability of the dataset. Then, following the same steps, we reach the formula for the second principal component:

$$Index_2 = 0.51 f_{20}$$

The regression results are shown in the second column of Table 3.6. The positive coefficient on years of schooling suggests that with more years of schooling, people are more likely to have negative attitude towards religion. However, this result is not statistically significant.

3.6 Robustness Checks

One concern of the first stage result is that the treatment indicator only reflects a time trend towards more schooling for younger cohorts. We test for the effects of placebo laws. If the instrument is valid, we should find no effects for any years earlier than the

---

Q3"People should not have any religious belief; instead people should only believe Science. 1. Agree; 2. Disagree."

year when the law was implemented.[32] Figure 3.4 plots the coefficients and 95% confidence intervals for the treated group using different years as the placebo effective time for the law. This figure displays the estimated coefficient of the treated group using data from before, during, and after the actual year of policy implementation. The point estimate at x=0 corresponds to the first stage estimation coefficient in Table 3.3. The point estimate at x=-1 is the placebo test result assuming the policy was implemented one year prior to the actual year. If 95% confidence interval contains 0, then there is no effect on years of schooling. We find that all years prior to the policy year have no impact on the years of schooling. Figure 3.4 suggests that the compulsory school law is a valid IV.

To ensure that the first stage results are not random, we performed a placebo test. To do so, we randomly assigned the policy implementation year of each province from a set of 12 years (5 years prior to and 6 years after the actual policy year), and use the placebo policy implementation year to define the treated and control groups. Then we run the first stage regression. The process is repeated for 500 times and the estimated coefficient of the treatment effect is presented in Appendix Figure 3.1. Appendix Figure 3.1 plots the distribution of the estimated effect of the policy on educational attainment based on these 500 iterations and the vertical line depicts the coefficient estimated using the actual policy implementation time. The actual value of the treatment effect is bigger than the 97th percentile of this distribution. This result suggests that the probability that the first stage result is generated purely by chance is less than 3 percent.

[32] Since the law states compulsory education as a long-term goal instead of an immediate change, we expect that placebo years later than the policy year to have larger effect. Local government may take years to implement this education policy. Economically developed areas were expected to make it universal faster than economically underdeveloped areas.

3.7 Conclusion

In this paper, we make use of the compulsory school law in China to examine how educational attainment affects religious beliefs over the long term. It is the first paper examining religion and education in a large developing country. We find that each additional year of schooling reduces the probability of an individual having religious beliefs by 8%. This effect is twice as high as the effect found in developed countries (Hungerman 2014).

We have also examined the relationship between education and religion by constructing various measures of religious beliefs, including whether people believe in any supernatural forces, whether people think religion is important or beneficial, and whether people wear religious objects or participate in any religious activities. Our new measures provide more accurate descriptions of beliefs and spiritual life in China. Many of our findings suggest a negative causal effect of education on religion.

We have also studied the effects of education on different religions. We find that the observed overall effect is largely attributed to the effects on western and local religion, whereas education has insignificant impacts on Buddhism. Our findings provide evidence of secularization in a developing country. The estimates also raise the question of why education might lower religiosity. One possible explanation is the science education and Atheism that are taught in junior high schools in China. However, due to the small sample size, we are not able to examine the impact of education on some religious groups with highly educated populations, such as Confucianism.

The impact of education on religion can also come from peer effects at school.

For example, compared to the individuals in the control group, individuals in the treated group may have more peers at school who do not believe in religion. Thus, they are less likely to be religious. Therefore our result may be a mixture of the direct effect of education (internal effect of education) and the peer effects (external effect of education). We plan to further our analysis to differentiate the internal and external effects.

Table 3.1- Summary Table

	Control	Treated	Total	Difference
Independent variables:				
Year of schooling	10.51	11.90	11.27	-1.400***
	(3.102)	(2.835)	(3.040)	
age	40.42	25.37	32.25	15.12***
	(3.391)	(4.662)	(8.558)	
age^2	1644.9	665.4	1113.3	983.9***
	(273.4)	(234.1)	(549.6)	
female	0.510	0.540	0.527	-0.019
	(0.500)	(0.498)	(0.499)	
married	0.953	0.519	0.717	0.435***
	(0.212)	(0.500)	(0.450)	
father's religion	0.0972	0.109	0.104	-0.013
	(0.296)	(0.312)	(0.305)	
mother's religion	0.136	0.166	0.152	-0.030**
	(0.343)	(0.372)	(0.359)	
minority	0.0396	0.0476	0.0439	-0.004
	(0.195)	(0.213)	(0.205)	
city	0.368	0.313	0.338	0.0445**
	(0.482)	(0.464)	(0.473)	
town	0.168	0.204	0.187	-0.026*
	(0.374)	(0.403)	(0.390)	
Outcome variables:				
Religion	0.211	0.207	0.209	0.002
	(0.408)	(0.405)	(0.406)	
Buddhism	0.174	0.169	0.171	0.003
	(0.379)	(0.375)	(0.377)	
Western religion	0.0299	0.0320	0.0310	-0.000
	(0.170)	(0.176)	(0.173)	
Western belief	0.125	0.139	0.133	-0.013
	(0.331)	(0.346)	(0.339)	
Local belief	0.365	0.364	0.364	0.005
	(0.482)	(0.481)	(0.481)	
N	2,019	2,394	4,413	

Notes: This table presents the summary statistics for the key variables by treated and control group. Standard deviations are in parentheses.

Table 3.2- First Stage Regression

	full sample	male subsample	female subsample	urban subsample	rural subsample
treated	0.835***	0.675*	1.062***	1.275***	0.612*
	(0.204)	(0.296)	(0.275)	(0.317)	(0.261)
age	0.338***	0.283***	0.475***	0.525***	0.272***
	(0.049)	(0.070)	(0.069)	(0.071)	(0.066)
age^2	-0.005***	-0.005***	-0.007***	-0.007***	-0.005***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
female	-0.607***			-0.157	-0.854***
	(0.088)			(0.134)	(0.113)
married	-0.594***	0.079	-1.517***	-0.936***	-0.433*
	(0.158)	(0.244)	(0.187)	(0.245)	(0.201)
father's religion	-0.517*	-0.382	-0.666*	-0.436	-0.485
	(0.228)	(0.384)	(0.277)	(0.446)	(0.263)
mother's religion	0.744***	0.647*	0.867***	0.575	0.745***
	(0.183)	(0.318)	(0.220)	(0.336)	(0.219)
minority	-0.119	-0.180	0.043	0.044	-0.120
	(0.242)	(0.347)	(0.339)	(0.386)	(0.315)
city	2.290***	2.054***	2.563***		
	(0.109)	(0.161)	(0.143)		
town	1.495***	1.483***	1.511***		1.529***
	(0.125)	(0.191)	(0.163)		(0.128)
Mean of Y	11.439	11.754	11.153	12.700	10.740
N	4413	2096	2317	1476	2937

Notes: This table presents estimates of the first stage results from Equation (2). *Treated* is an indicator for students who are younger than 13-year-old when the policy is implemented. The dependent variable is years of schooling. Standard errors are in parentheses.

***Significant at 1 percent level

**Significant at 5 percent level

*Significant at 10 percent level

Table 3.3- Second Stage Regression

	Model 1			Model 2		
	OLS	First stage	Second stage	OLS	First stage	Second stage
treated		0.827***			0.813***	
		(0.206)			(0.207)	
school	-0.004		-0.081*	-0.003		-0.083*
	(0.003)		(0.040)	(0.003)		(0.041)
age	0.010	0.342***	0.037*	0.007	0.344***	
	(0.008)	(0.049)	(0.017)	(0.009)	(0.060)	
age^2	-0.000	-0.006***	-0.001*	-0.000	-0.006***	-0.001*
	(0.000)	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)
female	0.034*	-0.597***	-0.013	0.034*	-0.582***	-0.014
	(0.013)	(0.089)	(0.029)	(0.013)	(0.089)	(0.029)
married	-0.024	-0.604***	-0.068*	-0.024	-0.609***	-0.069*
	(0.023)	(0.158)	(0.033)	(0.023)	(0.157)	(0.033)
father's religion	0.307***	-0.519*	0.267***	0.308***	-0.518*	0.266***
	(0.042)	(0.231)	(0.050)	(0.042)	(0.237)	(0.051)
mother's religion	0.274***	0.761***	0.333***	0.272***	0.758***	0.334***
	(0.036)	(0.185)	(0.049)	(0.036)	(0.187)	(0.050)
minority	0.003	-0.095	-0.003	0.002	-0.052	-0.000
	(0.034)	(0.247)	(0.041)	(0.034)	(0.248)	(0.042)
city	0.039*	2.274***	0.216*	0.039*	2.246***	0.218*
	(0.016)	(0.109)	(0.092)	(0.016)	(0.110)	(0.094)
town	0.028	1.474***	0.142*	0.026	1.474***	0.144*
	(0.019)	(0.126)	(0.063)	(0.019)	(0.126)	(0.065)
Mean of Y	0.215	11.442	0.215	0.215	11.442	0.215
f-stat		16.116			15.422	
N	4329			4329		

Notes: This table presents estimates of the second stage results from Equation (1). *Treated* is an indicator for students who are younger than 13-year-old when the policy is implemented. The dependent variable is a binary indicator, which equals 1 if the individual is reported to believe in any religion. In model 1, control variables includes province fixed effect, while in model 2, we control for province specific time trends. Standard errors are in parentheses.
***Significant at 1 percent level
**Significant at 5 percent level
*Significant at 10 percent level

Table 3.4-The Impact of education on Other Outcome Variables

	(1) Buddhism	(2) Western believe	(3) Religion importance	(4) Western supernatural belief	(5) Local supernatural belief
First stage:					
treated	0.827***	0.827***	0.800***	0.835***	0.835***
	(0.206)	(0.206)	(0.211)	(0.204)	(0.204)
Second stage:					
school	-0.033	-0.040*	-0.029	-0.125**	-0.102*
	(0.035)	(0.018)	(0.073)	(0.045)	(0.049)
Mean of Y	0.178	0.030	0.507	0.142	0.386
f-stat	16.116	16.116	14.438	16.820	16.820
N	4329	4329	4096	4413	4413

Notes: This table presents estimates of the second stage results from Equation (1). *Treated* is an indicator for students who are younger than 13-year-old when the policy is implemented. Each column represents a regression. The dependent variable corresponds to the column name. The control variables includes age, gender, marital status, minority status, parents' beliefs, location and province fixed effect. Standard errors are in parentheses.

***Significant at 1 percent level

**Significant at 5 percent level

*Significant at 10 percent level

Table 3.5- The Eigenvalues from PCA

Component	Eigenvalue	Difference	Proportion	Cumulative
Panel A:				
1	1.85		0.37	0.37
2	0.93	0.92	0.18	0.56
3	0.85	0.08	0.17	0.72
4	0.74	0.11	0.14	0.87
5	0.62	0.12	0.12	1.00
Panel B:				
1	1.52		0.51	0.51
2	0.99	0.53	0.33	0.84
3	0.48	0.52	0.16	1.00

Notes: This table presents the eigenvalues from PCA. In panel A, these components are used to construct the index religiosity and in panel B, these are used to construct attitude towards religion.

Table 3.6- The Impact of Education on Religiosity and Attitude

	(1) Religiosity	(2) Attitude
First stage:		
treated	0.864***	0.839**
	(0.217)	(0.261)
Second stage:		
school	-0.098	-0.112
	(0.089)	(0.119)
age	-0.011	-0.072
	(0.038)	(0.062)
age^2	-0.000	0.001
	(0.001)	(0.001)
female	-0.045	0.081
	(0.070)	(0.093)
married	-0.025	0.229**
	(0.076)	(0.112)
father's religion	0.302***	-0.138
	(0.125)	(0.164)
mother's religion	0.359***	-0.241*
	(0.121)	(0.147)
Mean of Y	-0.02	-0.08
f-stat	15.799	10.360
N	3723	2774

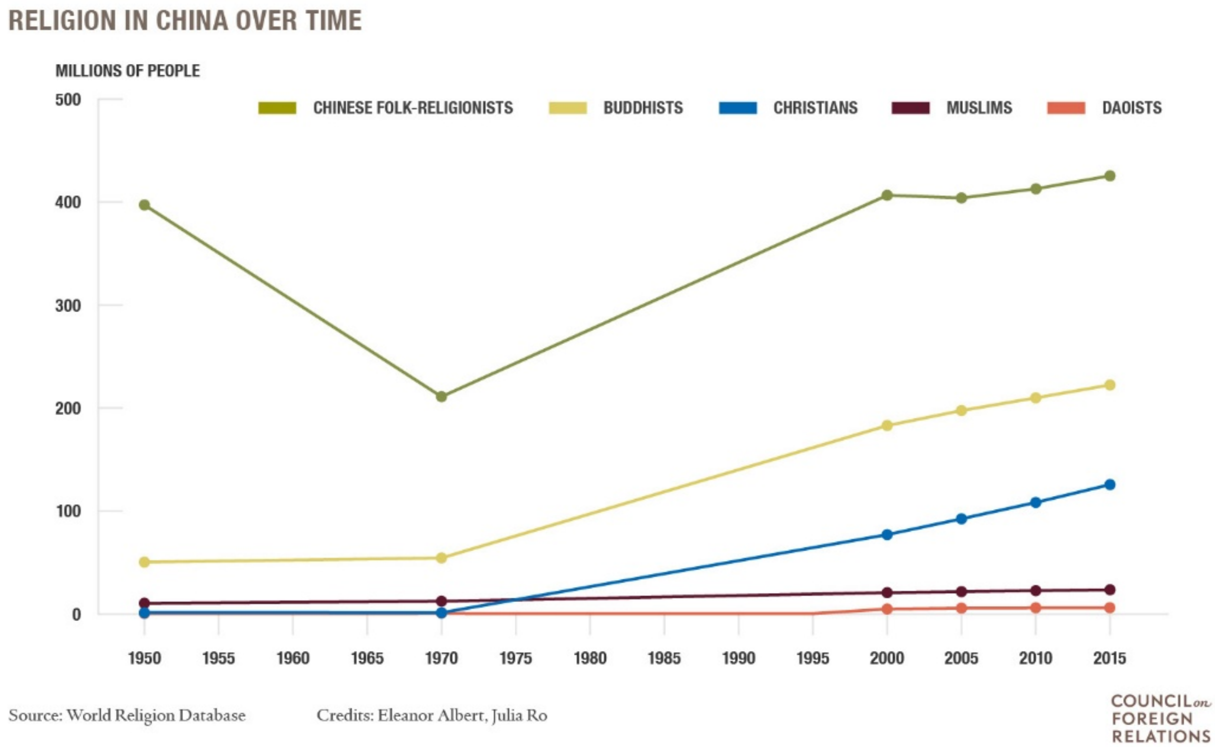
Notes: This table presents estimates of the second stage results from Equation (1). *Treated* is an indicator for students who are younger than 13-year-old when the policy is implemented. Each column represents a regression. The dependent variable corresponds to the column name. The control variables includes age, gender, marital status, minority status, parents' beliefs, location and province fixed effect. Standard errors are in parentheses.

***Significant at 1 percent level

**Significant at 5 percent level

*Significant at 10 percent level

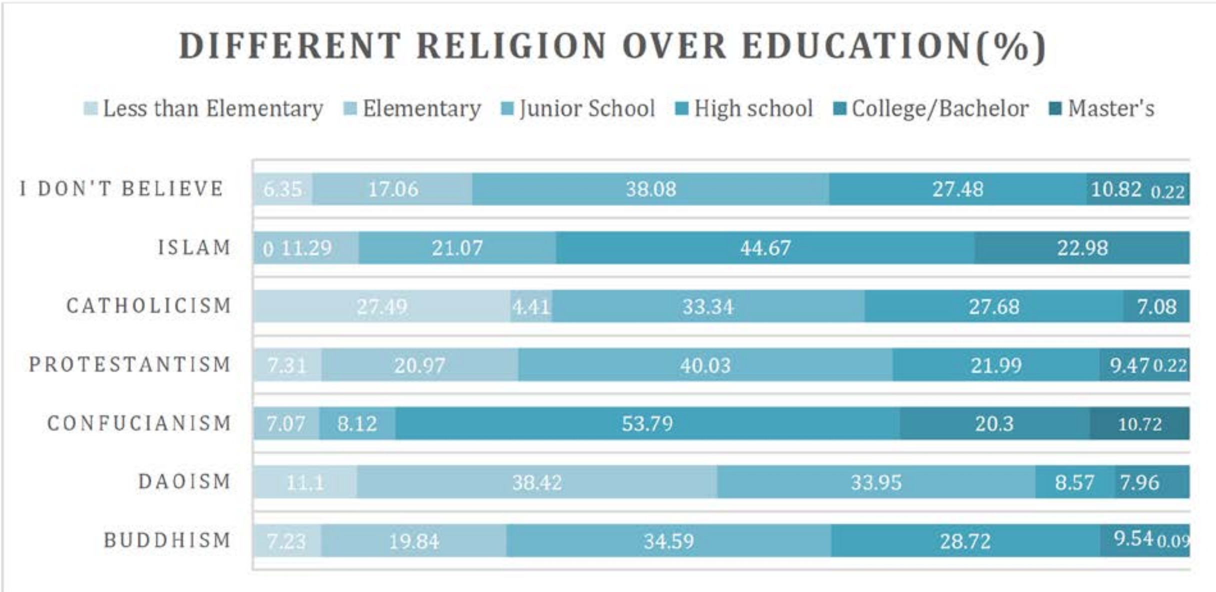
Figure 3.1- Religion in China over Time



Notes: Figure 3.1 presents the total number of people believe in five categories of religions over 1950-2015.

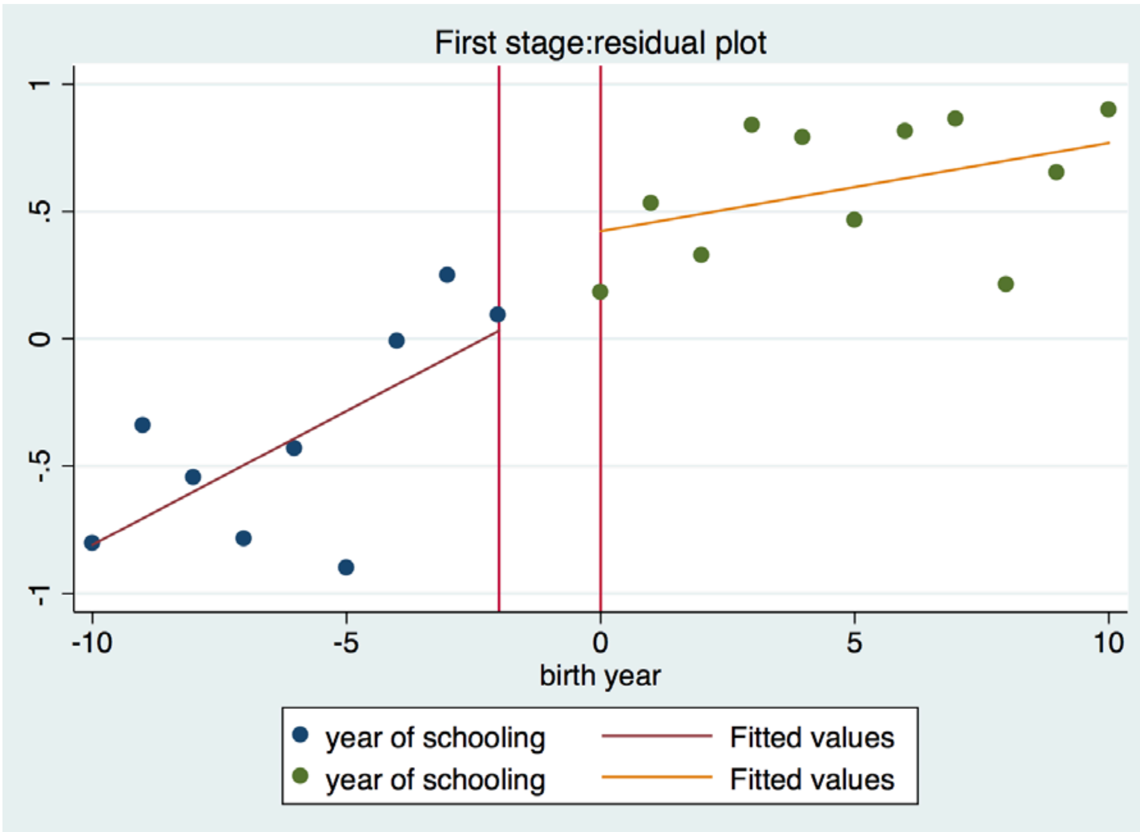
Figure 3.1. Religion in China over time. Reprinted from *Council on Foreign Relations*, by E.Albert and , J.Ro, 2015, retrieved from http://www.cfr.org/china/religion-china/p16272. Copyright[2015] by by E.Albert and , J.Ro..

Figure 3.2-Distribution of Religion over Education Groups



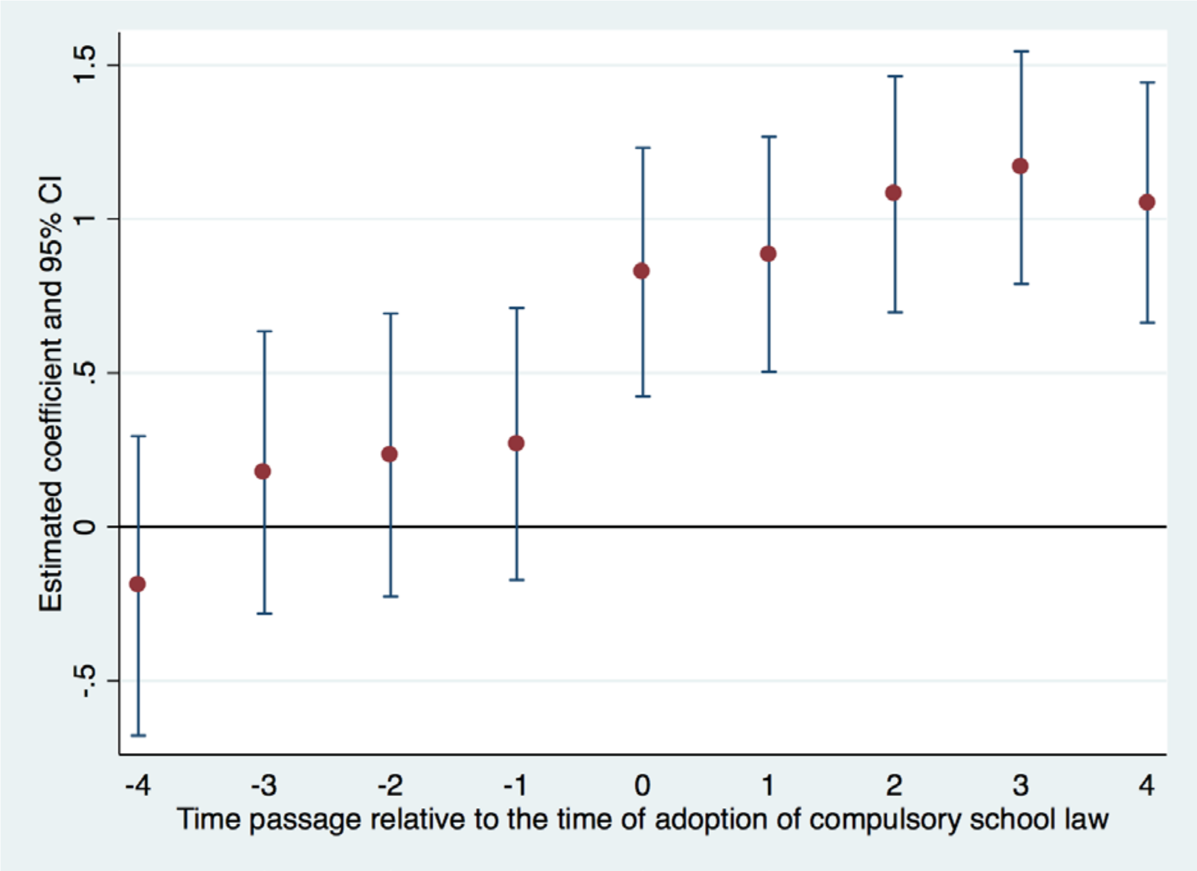
Notes: Figure 3.2 presents the distribution of education group of each category of religion.

Figure 3.3-First Stage Residual Plot



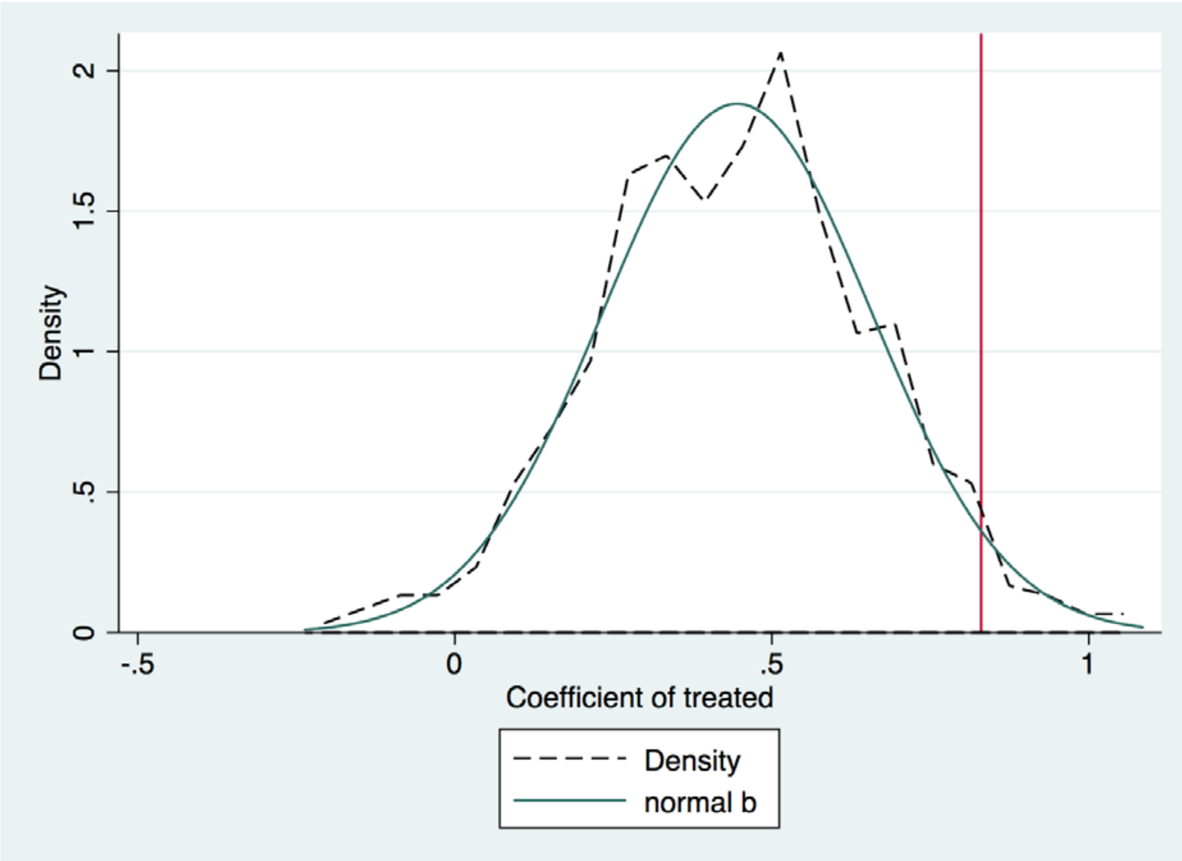
Notes: Figure 3.3 presents the relationship between the residual year of schooling and the distance to the implementation year of this compulsory school law. The x-axis is the difference between the year of age 13 and the implementation of the policy, and y-axis is the residual year of education. Cutoffs for treated group is represented by the vertical lines.

Figure 3.4- Robustness Check



Notes: Figure 3.4 plots the estimated results with 95% confidence intervals for placebo treatment year for Equation (2). The treated group is defined by different years as placebo time for the law. The x-axis is the difference between the placebo year and the actual year for policy implementation, and y-axis is estimated coefficients on treated indicator.

Appendix Figure 3.1-Placebo Test



Notes: Appendix figure 3.1 presents the distribution of the estimated effect of the placebo policy on years of schooling using 500 iterations. The vertical line depicts the coefficient estimated using the actual policy time.

Bibliography

1. Ammermuller, A., & Dolton, P. (2006). Pupil–teacher gender interaction effects on scholastic outcomes in England and the USA. *ZEW – Centre for European Economic Research: Discussion Paper No. 06-060.* *http://dx.doi.org/10.2139/ssrn.927689*.

2. Arcidiacono, P., Aucejo, E. M., & Hotz, V. J. (2016). University differences in the graduation of minorities in STEM fields: Evidence from California. *The American Economic Review*, *106*(3), 525-562.

3. Bar, T., Kadiyali, V., & Zussman, A. (2009). Grade information and grade inflation: The Cornell experiment. *The Journal of Economic Perspectives*, *23*(3), 93-108.

4. Barnett, M. D., Sonnert, G., & Sadler, P. M. (2012). More like us: The effect of immigrant generation on college success in mathematics. *International Migration Review*, *46*(4), 891-918.

5. Betts, J. R., & Shkolnik, J. L. (2000). The effects of ability grouping on student achievement and resource allocation in secondary schools. *Economics of Education Review*, *19*(1), 1-15.

6. Betts, J. (2011). The economics of tracking in education. *Handbook of the Economics of Education*, *3*(341-381), 4.

7. Bleakley, H., & Chin, A. (2004). Language skills and earnings: Evidence from childhood immigrants. *Review of Economics and statistics*, *86*(2), 481-496.

8. Borjas, G. J. (2005). The labor market impact of high-skill immigration (No. w11217). *National Bureau of Economic Research*.

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

10. Brown, S., & Taylor, K. (2007). Religion and education: evidence from the National Child Development Study. *Journal of Economic Behavior & Organization*, *63*(3), 439-460.

11. Card, D. (1999). The causal effect of education on earnings. *Handbook of Labor Economics*, *3*, 1801-1863.

12. Card, D. (2005) Is the new immigration really so bad?, *Economic Journal*, *115*, F300-F323.

13. Carnevale, A. P., Smith, N., & Melton, M. (2011). STEM: Science Technology Engineering Mathematics. *Georgetown University Center on Education and the Workforce*.

14. Carrell, S. E., Page, M. E., & West, J. E. (2010). Sex and science: How professor gender perpetuates the gender gap. *Quarterly Journal of Economics*, *125*(3).

15. Carrington, B., Tymms, P., & Merrell, C. (2008). Role models, school improvement and the 'gender gap' do men bring out the best in boys and women the best in girls? British *Educational Research Journal*, 34(3), 315–327 http://dx.doi.org/10.1080/01411920701532202

16. Cesur, R., & Mocan, N. H. (2013). Does secular education impact religiosity, electoral participation and the propensity to vote for Islamic parties? Evidence from an education reform in a Muslim country (No. w19769). *National Bureau of Economic Research.*

17. Chmielewski, A. K., Dumont, H., & Trautwein, U. (2013). Tracking effects depend on tracking type: An international comparison of students' mathematics self-concept. *American Educational Research Journal*, *50*(5), 925-957.

18. Cohen, M. L. (1992). Religion in a state society: China. *Asia: Case Studies in the Social Sciences*, 17-31.

19. Crisp, G., Nora, A., & Taggart, A. (2009). Student characteristics, pre-college, college, and environmental factors as predictors of majoring in and earning a STEM degree: An analysis of students attending a Hispanic serving institution. *American Educational Research Journal*, *46*(4), 924-942.

20. Dai, D. Y., & Rinn, A. N. (2008). The big-fish-little-pond effect: What do we know and where do we go from here? *Educational Psychology Review*, *20*(3), 283-317.

21. Dee, T. S. (2007). Teachers and the gender gaps in student achievement. *Journal of Human Resources* 42:528–54

22. Dee, T., & Lan, X. (2015). The achievement and course-taking effects of magnet schools: Regression-discontinuity evidence from urban China. *Economics of Education Review*, *47*, 128-142.

23. Ding, W., & Lehrer, S. F. (2007). Do peers affect student achievement in China's secondary schools? *The Review of Economics and Statistics*, *89*(2), 300-312.

24. Du, Yufang(2010). Current situation and development trend of religions in contemporary China. *Journal of the Central Institute of Socialism* 5 (2010): 77-83.

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

26. Ellison, G., & Swanson, A. (2010). The gender gap in secondary school mathematics at high achievement levels: Evidence from the American Mathematics Competitions. *The Journal of Economic Perspectives*, *24*(2), 109-128.

27. Elsner, B., & Isphording, I. E. (2015a). A big fish in a small pond: Ability rank and human capital investment.

28. Elsner, B., & Isphording, I. E. (2015b). Rank, sex, drugs, and crime.

29. Fang, H., Eggleston, K. N., Rizzo, J. A., Rozelle, S., & Zeckhauser, R. J. (2012). The returns to education in China: Evidence from the 1986 compulsory education law (No. w18189). *National Bureau of Economic Research*.

30. Figlio, D &Page, M. (2002). School choice and the distributional effects of ability tracking: Does separation increase inequality? *Journal of Urban Economics* 51, no. 3:497-514

31. Fischer, S. (2017). The downside of good peers: How classroom composition differentially affects men's and women's STEM persistence. *Labour Economics*.

32. Francis, B., Skelton, C., Carrington, B., Hutchings, M., Read, B., & Hall, I. (2006). A perfect match? Pupils' and teachers' views of the impact of matching and learners by gender. *Research Papers in Education*, 23(1), 21–36 http://dx.doi.org/10.1080/02671520701692510

33. Fryer, R. G., & Levitt, S. D. (2010). An empirical analysis of the gender gap in mathematics. *American Economic Journal: Applied Economics*, *2*(2), 210-240.

34. Gelman, A., & Imbens, G. (2014). Why high-order polynomials should not be used in regression discontinuity designs (No. w20405). *National Bureau of Economic Research.*

35. Glaeser, E. L., & Sacerdote, B. I. (2008). Education and religion. *Journal of Human Capital*, *2*(2), 188-215.

36. Griffith, A. L. (2010). Persistence of women and minorities in STEM field majors: Is it the school that matters?. *Economics of Education Review*, *29*(6), 911-922.

37. Gruber, J. H. (2005). Religious market structure, religious participation, and outcomes: Is religion good for you?. *The BE Journal of Economic Analysis & Policy*, *5*(1).

38. Gulesci, S., & Meyersson, E. (2012). For the love of the republic: education, religion, and empowerment. *Working paper*.

39. Hanushek, E. A., & Woessmann, L. (2012). Do better schools lead to more growth? Cognitive skills, economic outcomes, and causation. *Journal of Economic Growth*, *17*(4), 267-321.

40. Ho, D. E., & Kelman, M. G. (2014). Does class size affect the gender gap? A natural experiment in law. *The Journal of Legal Studies*, *43*(2), 291-321.

41. Hoffer. (1992). Middle School Ability Grouping and Student Achievement in Science and Mathematics. *Educational Evaluation and Policy Analysis 14*, no. 3 (1992): 205-27.

42. Hume, D. 1757. The natural history of religion, edited by J.C.A. Gaskin. *Oxford:Oxford University Press*

43. Hungerman, D. M. (2010) "Does religious proscription causes people to act differently? Evidence from a theory- based test," *Working paper.*

44. Hungerman, D. M. (2014). The effect of education on religion: Evidence from compulsory schooling laws. *Journal of Economic Behavior & Organization*,*104*, 52-63.

45. Hyde, J. S., Lindberg, S. M., Linn, M. C., Ellis, A. B., & Williams, C. C. (2008). Gender similarities characterize math performance. *Science*, *321*(5888), 494-495.

46. Imbens, G. W., & Lemieux, T. (2008). Regression discontinuity designs: A guide to practice. *Journal of econometrics*, *142*(2), 615-635.

47. Lam, K (2011,Nov 1) *Curriculum Reform in Chinese Secondary Education* Retrieved from http://wenr.wes.org/2011/11/wenr-novemberdecember-2011-practical-information-curriculum-reform-in-chinese-secondary-education

48. Lavy, V., & Sand, E. (2015). On the origins of gender human capital gaps: short and long term consequences of teachers' stereotypical biases (No.w20909). *National Bureau of Economic Research.*

49. Lim, J., & Meer, J. (2015). The impact of teacher-student gender matches: Random assignment evidence from South Korea (No. w21407). *National Bureau of Economic Research.*

50. Luppino, M., & Sander, R. H. (2013). College major competitiveness and attrition from the sciences.

51. Machin, S., & Murphy, R. (2014). Overseas students: the impact on domestic student numbers (No. 426). *Centre for Economic Performance, LSE.*

52. Ma, M., & Shi, X. (2014). Magnet classes and educational performance: Evidence from China. *Economic Development and Cultural Change*, *62*(3), 537-566.

53. Manski, C. F. (1993). Identification of endogenous social effects: The reflection problem. *The review of economic studies, 60*(3), 531-542.

54. Manski, C. F. (2000). Economic analysis of social interactions (No. w7580). *National bureau of economic research.*

55. Marsh, H. W. (1987). The big-fish-little-pond effect on academic self-concept. *Journal of educational psychology*, *79*(3), 280.

56. Marsh, H. W., Kong, C. K., & Hau, K. T. (2000). Longitudinal multilevel models of the big-fish-little-pond effect on academic self-concept: counterbalancing contrast and reflected-glory effects in Hong Kong schools. *Journal of personality and social psychology*, *78*(2), 337.

57. Marsh, H. W., Trautwein, U., Lüdtke, O., Köller, O., & Baumert, J. (2005). Academic self‐ concept, interest, grades, and standardized test scores: Reciprocal effects models of causal ordering. *Child development*, *76*(2), 397-416.

58. McCleary, R. M., & Barro, R. J. (2006). Religion and political economy in an international panel. *Journal for the Scientific Study of Religion*, *45*(2), 149-175.

59. Meyersson, E. (2009). Islamic rule and the emancipation of the poor and pious. *Working Paper, IIES Stockholm University*.

60. Murphy, R., Weinhardt, F. (2014). Top of the class: The importance of rank position. *http://personal.lse.ac.uk/murphyrj/_private/topofclass.pdf*

61. Oreopoulos, P., & Salvanes, K. G. (2009). How large are returns to schooling? Hint: Money isn't everything (No. w15339). *National Bureau of Economic Research.*

62. Orrenius, P., & Zavodny, M. (2013). Does immigration affect whether US natives major in a STEM field? *Journal of Labor Economics*.

63. Osafo, J., Knizek, B. L., Akotia, C. S., & Hjelmeland, H. (2013). Influence of religious factors on attitudes towards suicidal behaviour in Ghana. *Journal of Religion and Health*, *52*(2), 488-504.

64. Ost, B. (2010). The role of peers and grades in determining major persistence in the sciences. *Economics of Education Review*, *29*(6), 923-934.

65. Powdthavee, N. (2010). Does education reduce the risk of hypertension? Estimating the biomarker effect of compulsory schooling in England. *Journal of Human Capital*, *4*(2), 173-202.

66. Rasmussen, C., & Ellis, J. (2013). Who is switching out of calculus and why. In *Proceedings of the 37th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 4, pp. 73-80).

67. Rawlings, S. (2014). Parental education and child health: Evidence from an education reform in China.

68. Rees, Brewer, and Argys(2000). How should we measure the effect of ability grouping on student performance? *Economics of Education Review 19,* no. 1: 17-20.

69. Silles, M. A. (2009). The causal effect of education on health: Evidence from the United Kingdom. *Economics of Education Review*, *28*(1), 122-128.

70. Slavin, Robert E. (1987). Ability grouping and student achievement in elementary schools: A best-evidence synthesis. *Review of Educational Research 57,* no. 3: 293-336.

71. Tipping, M. E., & Bishop, C. M. (1999). Probabilistic principal component analysis. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *61*(3), 611-622.

72. Trautwein, U., Lüdtke, O., Roberts, B. W., Schnyder, I., & Niggli, A. (2009). Different forces, same consequence: conscientiousness and competence beliefs are independent predictors of academic effort and achievement. *Journal of personality and social psychology*, *97*(6), 1115.

73. Valentine, J. C., DuBois, D. L., & Cooper, H. (2004). The relation between self-beliefs and academic achievement: A meta-analytic review. *Educational Psychologist*, *39*(2), 111-133.

74. Vardardottir, A. (2013). Peer effects and academic achievement: A regression discontinuity approach. *Economics of Education Review*, *36*, 108-121.

75. Xie, S., & Mo, T. (2014). The impact of education on health in China. *China Economic Review*, *29*, 1-18.

76. Zhou, Li. "Native students and the gains from exporting higher education: Evidence from Australia." (2011).