Essays on the Economics of Education

Steven T. Simpson

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

COLUMBIA UNIVERSITY

2013

ABSTRACT

Essays on the Economics of Education

Steven T. Simpson

Post-secondary education is becoming increasingly more common for students around the world. As quantity of education increases, it becomes less of a distinguishing factor to be simply a college graduate. For those who want to stand out, the quality aspects of education become more salient. Moreover, as this expansion happens in the number of colleges and college students, it becomes less common for governments to generously fund the college education of a lucky few. In addition, the cost to colleges to provide an education is also increasing. Taken together, simply as a measure of cost-comparison, choosing between colleges based on the potential quality-for-money is also an important reason for college quality's increasing salience. College quality matters, and this dissertation endeavors to show how and to what extent.

The following three separate chapters estimate the returns to different forms of college quality. There has been an extensive literature that shows, in general, that more schooling is better. These chapters seek to shift the margin of analysis from the extensive margin of quantity to the intensive margin of quality. Thus, I ask the question: is better schooling better or, to put it another way, how much better is better schooling?

In the first chapter, I estimate the returns to college quality, operationalized mainly through peer quality, using a regression discontinuity design and exploiting the two separate rounds (early and regular) of college admissions in Taiwan. In the second chapter, focusing on college prestige, I again use a regression discontinuity design to estimate the returns to scoring just above (vs. just below) the admissions cutoff for the lowest-ranked national college. The theory of action is that national colleges are uniformly more desirable than

private colleges (excluding a few elite private colleges), if for no other reason than that their tuitions are subsidized by the government and thus much lower for the individual. The final chapter looks at a set of 11 “colleges” that had already been meeting the minimum requirements for being labeled a “university” (an important distinction in Taiwan’s system), but for bureacratic reasons had not been allowed to change their label/rank until a policy change in 1997. Treating this policy change as a natural experiment, I use a difference-in-differences framework to show that cohorts entering these newly upgraded 11 universities earn statistically significantly more than cohorts entering prior to the change at the same colleges.

A consistent picture emerges out of these three papers: college quality matters on several dimensions. These chapters are set apart from other papers in the literature by the causal interpretation given to both choice of college AND choice of college major. My estimates show that those who attend higher quality colleges, within the same college major, end up earning between one-tenth to one-fifth of a standard deviation (σ) more in their first year of employment after graduating. Peer quality, college prestige, and college reputation all appear to provide a return. But choice of college major appears to be one of the most important dimensions through which college quality operates, with the Science-track college majors receiving most of those returns to quality.

Table of Contents

List of Figures

List of Tables

Acknowledgments

This dissertation has benefited from discussions with and comments from Jin-Tan Liu, Eric Verhoogen, Bernard Salanie, Todd Kumler, Giovanni Paci, Vicente Garcia-Moveno, Thomas Bailey, and from the participants of the Development and Applied Micro Theory Colloquia at Columbia University.

I especially thank my advisors who have provided much guidance and support during my dissertation. Miguel Urquiola provided helpful feedback, support, time, and patience, all of which have made me into a better economist. Judith Scott-Clayton has been critical to my success here at Columbia. Serena Ng has been supportive throughout this program.

This dissertation would not have been possible without the trust of Jin-Tan Liu in sharing his data. Thank you for taking a chance on my ideas.

I thank my classmates and friends who have supported me through the ups and downs of graduate school and living in New York City. You have made the five years so enjoyable. I am grateful to my family, who has always been supportive and loving in whatever I have pursued. To Lish, thank you so much for your love, support, patience, and encouragement. I could not have done this without you.

Most of all, I thank and praise Jesus, the true example of the brightest intellect and purest goodness selflessly spent on educating the lost and empowering the weak—indeed, completely saving the helpless through his total sacrifice.

to Lish

my Beloved and my Chum,

to Makeila and Livia

my delightful daughters and my greatest Calling,

and to our parents

who have been examples and enablers

Chapter 1

Returns to College Quality: Colleges, Classmates, and College Majors

College quality matters. Three global trends highlight college quality's rise to prominence in individual decisions and policy discussions regarding higher education. First, the UNESCO World Conference on Higher Education reports that between 2000 and 2007 the gross enrollment rates have changed dramatically in some regions of the world. For example, in Latin America the increase has been by one-third; in Central and Eastern Europe, it has been by one-half; and in Asia it has been a full doubling (Altbach et al., 2009). Importantly, this increase has been for both males and females alike, even with females out-pacing males in many contexts, including the US (Becker et al., 2010; Goastellec, 2008). With this high rate of expansion in the quantity of higher education, a relevant concern has been whether or not the quality of higher education can keep apace.

Second, past research suggests that a college education has been one of the best investments an individual could make in her future (Goldin and Katz, 2009). As individuals face a job market increasingly crowded with other college graduates, a college degree seems less the guarantee of 'the good life' it used to be. For example, employers may no longer be asking *whether* the job applicant went to college, but rather *where* the applicant went. Consequently, students have had to switch from competing on the extensive margin for more quantity of education to the intensive margin for higher quality of education. Relatedly, college selectivity has been one way graduates have sought to distinguish themselves. Bound et al. (2009) show that the number of universities that US high school students apply to has increased steadily and that the growth in applications has been disproportionately weighted towards selective institutions. Third, institutions of higher quality or greater selectivity cost more, either to the individual or to the general public; it matters then how much more quality the higher cost is purchasing.

The research on college quality thus far has found it difficult to precisely and convincingly estimate the importance of college quality to later-life outcomes. I contribute to the college quality literature by using data from Taiwan and a regression discontinuity (RD) design that compares outcomes for students scoring just above the admissions threshold versus those scoring just below. This comparison provides an attractive research design because students just above and below an admissions cutoff should be similar in expected outcomes; but because of where they fell in relation to the cutoff, they end up having different college experiences. Consequently, comparing later-life outcomes for these students should, on average, provide an estimate of the returns to the differences in the college quality they experienced.

Using this RD design, I find in my study that those scoring above the cutoff are exposed to higher quality colleges: these institutions have 0.04 σ (standard deviations) better average peer quality, are 0.10 σ more likely to be of national (more prestigious) rank, and are 0.12 σ more likely to be in the top-6 of elite institutions. These first-stage quality advantages correspond to a 3.1 log-point *reduction* in yearly tuition, or about 0.1 σ in subsidized tuition. Upon graduation, those students who scored above the cutoff earned 3.2 log-points (on a 2-digit log base) or 0.12 σ higher wages, which is equivalent to 11100 NTD or 5% more each month. These results are economically substantive, statistically significant, and robust to specification testing.

My research furthers the current 'college quality' literature in three areas: data quantity, data quality, and research design. Mine is the first RD paper to use a national universe of microdata for college applications, college admissions, and earnings — all merged on national IDs — to identify causal estimates of the returns to college quality. Because the data I use catalogs student admissions down to the academic department level, I can estimate these returns not only to the choice of college, but also to the choice of college major — something which had previously been the purview of structural estimation strategies (Arcidiacono, 2004).

Second, in terms of data quality, I have not only the universe of data for this admissions cohort but also the universe of discontinuities. My reduced-form results are based on an

RD framework with nearly 700 discontinuities generated by the Taiwanese government's centralized student allocation system.

Third, in terms of research design, the stacked RD I use allows me to approximate a local average treatment effect (LATE) at every admissions cutoff in the test score distribution. In this study, the LATE is the estimate of the causal effect of scoring above the cutoff for where the student applied: that is, for getting into a better college. The LATE is described in the economics literature as being 'local' since it applies only to those students around a particular cutoff, thus limiting its generalizability. When making a policy, the average treatment effect is the most preferred estimate, because it represents the population of interest, not just one of its subgroups. Therefore, having only a *local* average treatment effect means that the LATE falls short of the optimal policy measure (Imbens, 2010). However, by having a LATE at *every* cutoff in the test-score distribution and for *every* academic department in the country, I recover a broad degree of the generalizability needed for policymaking (Pop-Eleches and Urquiola, 2012). In addition, I am able to explore the heterogeneity in outcomes that is also very relevant to policymaking. These three aspects of my study allow me to convincingly and causally estimate the wage returns to college quality using every college-goer and every college in a single market (for both higher education and the labor force) — a first for the college quality literature.

Like other recent papers in the college quality literature (Black and Smith, 2006; Dale and Krueger, 2002, 2011), I measure college quality as a function of peer quality, proxied by the average college entrance exam score at that college. I also measure college quality using college prestige, proxied by a college's status as national rank (Chevalier and Conlon, 2003) and the amount of public subsidy a student receives (Van der Klaauw, 2002; Winston, 1999). In the body of the literature, other measures of college quality include endowments (Winston, 1999), faculty quality and faculty-student ratios (Saavedra, 2008; Winston, 1999), and research productivity (Aghion et al., 2010). Each of these papers shows the correlation between peer quality and other proxies of college quality, thus making my proxy a useful summary. Reduced-form outcomes to school quality, such as aggregate productivity in the workforce or economic growth (Hanushek and Woessmann, 2012; Lucas, 1988), however, are beyond the scope of this paper.

In addition to focusing on college quality, my paper contributes more broadly to the 'better schools' literature by estimating wage returns in the labor market — something past studies in the specific literature have been unable to show. This literature (Clark, 2007; Jackson, 2010; Pop-Eleches and Urquiola, 2012) has used RD designs and generally found that going to a better high school can lead to higher scores, though not more tries, on a college entrance exam. Despite the consensus most studies have on the benefits of better schools, Abdulkadiroğlu et al. (2011) provide a dissenting voice, showing that attending top exam schools in Boston and New York provides no additional benefit on a battery of state standardized exams and the SAT.

The remainder of the paper is organized as follows: Section 1 discusses the educational system in the Taiwan context; Section 2 details and summarizes the administrative data sets used; Section 3 motivates the use of RD estimation strategies; Section 4 discusses main results; Section 5 explores heterogeneity; and Section 6 concludes.

1.1 Literature and Context

One way to estimate the importance of college quality on future earnings is to conduct a survey using a random sampling of the relevant subpopulation of college-goers and asking each respondent where s/he went to college and how much his/her income is. Producing a summary table of these results would give an accurate description of the relationship between college quality and earnings. Moving beyond describing the relationship, one could estimate this correlation using a form of statistical analysis, such as regression. However, as several researchers have noted in the past, there are individual characteristics — sex, race, family background, to name a few — that could predict both the probability of attending a certain type and quality of college as well as the level of income that person achieves. Each of these characteristics would have been set prior to ever attending college. Thus, the earnings benefit observed may be better attributed to these factors rather than to college quality.

The correlation between these pre-existent characteristics has led many to try to account for them by including them as controls in regression models. A serious concern with this strategy, however, is the difficulty of determining how many relevant factors exist. Omitting any one of these factors will bias the results against finding the true estimate. Lee and Lemieux (2010) point out that even when a 'kitchen sink' regression includes all relevant individual characteristics, there is still no way to estimate the returns to college quality independent of these covariates, because in many cases there is no independent variation left: the covariates perfectly predict the outcomes.

Because of the shortcomings of the two above methods, researchers have recently been looking beyond statistical fixes like "selection on observables" and matching, and moving towards research designs similar to those used in experimental settings. Because randomized experiments are expensive and often come up against ethical objections, a set of natural (or quasi-)experimental methodologies have been designed to identify and use naturally occurring opportunities to approximate a randomized experiment so that cross-section data can still be used. The regression discontinuity design is considered to be one of the most convincing of these research designs (Angrist and Pischke, 2010).

The following key papers illustrate the evolution in the college quality literature over the past two decades towards the quasi-experimental approach in the economics literature. Dale and Krueger (2002) is one of the early papers in the natural experiments literature on school quality that sought to overcome selection problems between schools and students. Using the College and Beyond Survey data, they observe which colleges students applied to, were admitted at, and eventually attended. They find that compared to those who were accepted but chose not to go, those who attended the more selective institution experienced a significantly higher wage in their first year of employment. However, Dale and Krueger show that 5 years later, the advantage in earnings diminishes, which they interpret to mean that college quality does not matter. Dale and Krueger (2011) return to this analysis using both more and better administrative earnings data further out in time and a more recent larger set of students. Their conclusion is the same: controlling for average SAT score—a measure of pre-existing student ability—absorbs all the effects of attending a selective college.

Dale and Krueger (2002, 2011) contain two critical weaknesses. First they assume that random reasons led students who were accepted into a more selective institution to reject their offers there. This is unlikely. Second, despite being a novel data experiment, fundamentally Dale and Krueger are using a control strategy that depends on including many covariates (in the form of the average SAT score for every college that the student applied to) in an attempt to absorb all the influences of selection. By doing this, they may be unintentionally absorbing both good and bad variation, and thus their null finding may not be conclusive or surprising. MacLeod and Urquiola (2011) argue against Dale and Krueger's interpretation from another perspective. They believe that first-year earnings is exactly where we would expect college quality to play an important role. Employers use college quality to simplify the task of predicting a graduate's productivity before they have had a chance to observe it for themselves. The longer employers do observe the worker's productivity, the less they will need to rely on other signals, such as the average quality of the college the worker attended. This argument does not require that school quality work only as a signal of student ability. Indeed, MacLeod and Urquiola (2011)'s model of college reputation allows for both human capital and signaling effects.

Hoekstra (2009) is an early contribution to the RD literature on college quality — one that exemplifies the move away from elaborate control strategies towards cleanly identified natural experiments. Instead of the match and control strategies that Dale and Krueger use, he uses an RD design to compare the earnings outcomes for those who were marginally admitted and rejected at a flagship university in Texas. The quality of Hoekstra's application and earnings data is what makes his RD design possible. In addition to the detailed admissions records, Hoekstra uses administrative earnings data from IRS tax records for 10 years after the students graduated; this refinement of data is a significant improvement over studies that rely on less reliable, self-reported earnings of survey respondents. Hoekstra finds that those who were admitted to the selective college earned a significantly higher income than those who were not. Because of the higher standards imposed by the RD design, Hoekstra's conclusions are arguably more convincing than previous work using control strategies and 'kitchen-sink' regressions. However, a critical weakness of Hoekstra's design is that he does not know what happened to those who were not admitted in his sample. Importantly, he does not know for sure whether they even went to college or not. Thus, the earnings advantage he finds in his study could be simply the earnings return to college attendance, instead of college quality. There are reasons to believe, though, that those who were competing at the margin for admissions to the flagship college in Texas would have been over the admissions threshold at one of the non-flagship institutions, and very likely would have gone on to attend college.

Saavedra (2008)'s contribution is valuable at many levels. First, he has access to data on a large set of selective institutions; therefore, he observes which institutions students attended, even if they were not admitted into the selective college Saavedra's analysis focused on. Second, he uses an RD setup to estimate the returns to exceeding the test-score cutoff for admissions to a top school in Colombia. Finally, possibly one of the most important contribution of this study is that it is able to estimate the value-add of college quality on knowledge learned, using scores on subject-specific standardized college *exit* exams commonly taken by college graduates at selective institutions. Pulling together all this data into an RD design, Saavedra finds that those just admitted to the top college, compared to those just rejected, scored around a 0.2 σ higher on their exit exam and enjoyed first-year earnings about 35% higher. Rubinstein and Sekhri (2011) is the only other RD paper able to estimate the value-add of attending a more prestigious university. In most developing countries, public institutions are more prestigious than private ones, and thus are considered to produce better learning. Their findings are surprising in that they can find no value-add for attending the more prestigious public colleges, as measured by *exit* exam scores. In such a case, it would be valuable to explore whether wages differed as well for these graduation. Unfortunately, their analysis does not include wages.

For all of the papers described above, the analysis is dependent on measuring returns to attending a smaller set of selective institutions; this limits the external validity of their analysis. More than just excluding a majority of the colleges and college-hopefuls from the analysis, focusing the analysis on only the top schools and students may miss identifying the place in the overall distribution of quality where college quality has the largest effects. For instance, we may think that the largest return to college quality is not going to be experienced by moving from a selective institution to a more prestigious one, but rather moving from a non-selective institution into a selective one. In other words, it may, in fact, be that excluding those of the lowest ability from a student's peer group is more important than necessarily including those of highest ability. Because the market in Taiwan is completely selective, even at the lowest quality institutions, I am able to estimate the returns to college quality across the full range of the quality distribtion in Taiwan's market, which is a first for the literature. As detailed above, these gaps in the literature and the opposing views from its seminal papers both create space for new and varied evidence to add to the discussion.

1.1.1 General context in Taiwan

Taiwan's market for higher education is centrally controlled by the Ministry of Education (MOE), as is every level of education there. One of the historically distinctive characteristics of that centralized control is its constraint of supply at every level of education at both public and private schools. This control extends down to the number of students and teachers that are allowed in each college and academic department in any given year.

Early research on the returns to education in Taiwan focused on the 1968 expansion of compulsory schooling to include the 3 years of junior high school. With this reform, junior high school enrollment increased by 50% in the first year (Clark and Hsieh, 2000). Several studies have looked at the effect of this junior high school expansion on such varied outcomes as the labor market returns to education (Spohr, 2003), the decrease in wage inequality (Vere, 2005), the increase in female labor market participation (Tsai et al., 2009), the effect of maternal education on infant health (Chou et al., 2010), and the intergenerational transfer of human capital (Tsai et al., 2011).

A second historically distinctive characteristics of Taiwan's centralized control is its constraint on access. This part of the system has not received much attention in the literature. All admissions in Taiwan use nationally standardized tests and clear protocols for admissions. There are two rounds of admissions, early and regular. Both are described below. Figure 1 provides the schedule for early and regular admissions phases in the year.

1.1.2 Regular admissions

The MOE in Taiwan exercises full control over the public and private provision of higher education; this control is most evident in that no student can be admitted into any institution without going through the Taiwanese system of admissions. In order to make the highly competitive admissions process as meritocratic as possible, the MOE has centrally administrated all aspects of regular admissions. Approximately 93% of all students admitted to college in Taiwan do so through regular admissions. Students are not allowed to apply directly to the institution, nor can departments recruit students. Instead, regular admissions is based on students' scores from the Joint College Entrance Exam (hereafter "the July exams") — a once-a-year, national standardized battery of tests. After receiving the grades on the July exams, students must submit to the MOE a rank ordering of their preferences for attending specific departments within colleges.

Underpinning the regular admissions process is the Deferred Acceptance algorithm (hereafter "the algorithm") it uses. The algorithm considers a student for a slot at each of the college departments where she has applied. If her July exam scores are high enough to make her eligible for a slot at two or more departments, the algorithm gives her a slot in the department she prefers most, assigning the slots in the lesser preferred departments to others with lower scores. This process goes on until all slots are filled, leaving some students without a slot at any college department. The outcome of the matching is final. Departments are required to admit all students assigned to them by the algorithm; there are no post-admissions transfers between departments. The only way to be admitted into a department is to go through the MOE-defined system. Students, if they wish, are allowed to retake the exam the following year, but this is less common for already admitted students.

1.1.3 Early admissions

Culminating in the early 1990s, schools and parents of students complained that the July exams measure (and therefore encourage) only a limited number of student qualities that parents, schools, and even employers may care about: such as, charisma, articulateness, adaptability, or congeniality, to name a few. Therefore, to address this popular complaint, the MOE initiated an early admissions protocol to meet parental and institutional desires for greater comprehensiveness in the admissions process.

The set up of the early admissions process works as follows: starting in early February before regular admissions, any high school senior may sit for a set of 5 subject-specific tests, referred to as the Competency Exam (here after "the February exams").[1] Once students receive their scores for the February exams, they make their single choice of where to apply. Unlike regular admissions, early admissions allows students only *one* chance to apply *directly* to *one* specific department. Early admissions decisions are left to the department's admissions committee, who are closely supervised by the MOE.

One could be concerned that adding the early admissions process invalidates in some ways the efforts exerted to make regular admissions a fair matching based solely on merit. Indeed, similar complaints have been brought against early admissions in the US, with critics claiming it is a backdoor for elite applicants at elite institutions (Avery et al., 2004; Avery and Levin, 2010). However, several features of Taiwan's early admissions protocol protects the integrity of the system. First, the timing of the Taiwanese system is coordinated, and the same application requirements are used across all institutions. In contrast, US institutions regularly differ with each other on when early applications are due, what is required, and when students are notified about their admissions status (Avery et al., 2004). Moreover, across years and within each institution in the US, the requirements could change depending on which strategies these elite institutions use to compete with each other. Critics of the US early admissions system argue that the only individuals able to be competitve during early admissions in the US are those who have the resources to plan early, test well and often, and enroll independent of financial aid packages. In Taiwan, this is not the case. All tests, applications, and notifications take place at the same time and in the same way. Moreover,

[1] The February exams differ from the July exams in two ways: first, all 5 subject-tests are required for the February exams; second, the February exams test general knowledge of the type learned in high school, unlike the July exams which tests for track-specific abilities needed in college.

there is not a system of financial aid at the individual level. Either a student gets into a national college that is subsidized or a private one that is not.

Second, early admissions in Taiwan allows each student to apply early to only one department. This restriction prevents students (departments) from gaming the early admissions system by playing departments (students) off each other, as is done in the US (Avery et al., 2004). Third, departments can use only a share of their total slots during early admissions. In 2000 more than 90% of the departments used less than 25% of their total slots for early admissions, with the average being less than 10% used. This limitation is placed and enforced by the MOE, which monitors the departments closely throughout the early admissions season. Every year the share the department can use for early applicants is confirmed by the MOE before the application season starts, and the department's compliance is verified at the end of early admissions. The departments are required to announce to the MOE which early applicants they have offered early admissions to and which have accepted the offers. Students are not allowed to take the July exams if they are still holding an early admissions slot at a department.

This limitation on the share of total capacity used is critical not only to the efficiency of the mechanism design, but also to the construction of my RD design. For the efficiency of the admissions process, the MOE does not want a majority (or even a close minority) of the college students being admitted early through direct application, since it would significantly reduce the influence of their centralized control. For the construction of my regression discontinuity design, this binding constraint on the usable share of total department capacity generates a discontinuity in access for students applying early. The random process that causes one student to end up above versus below the discontinuity is the source of variation off which I identify the returns to college quality.

After the February exams, student applications are reviewed by departments, and shortlisted applicants are interviewed by the department's admissions committee. Before the July exams, early admissions offers are made to the students. At this point, students can choose to accept the admissions offer or reject it. Only about 5 percent of students who are offered a seat through early admissions reject the offer; these tend to be top-performers who are far away from the cutscore. In order to participate in the regular admissions, students who have been offered early admissions must give up their early admissions seat. This is because there are no wait-lists in Taiwan. All unfilled seats must be in play during the regular admissions process in order for the matching mechanism to be truly efficient. Those who compete in early admissions but fail must then compete in regular admissions in order to be admitted at all.

Since I observe in my data the admissions outcomes for both early and regular admissions, I am able to use both the bifurcation in timing between early and regular admissions as well as the discontinuity for each department in early admissions as one natural experiment to estimate the returns to college quality. To achieve a common measure of peer quality across both early and regular admissions, I use the regular admissions test scores in each department to generate a proxy for average peer quality. (See Section 2.1 for a description of how peer quality is calculated.) Since the majority of students enter any given department through regular admissions, I argue that the department average of July exam scores provides a useful proxy for the average peer a student was exposed to after being admitted. Thus, even for those students who were admitted early and never took the regular admissions exams, I impute the average of July exam scores. Because early admitted students eventually end up in the same departments with regular admissions students, the average July exam score is, indeed, a measure of the quality of the peers they share that department with.

1.2 Data

This paper uses a universe of governmental administrative data on birth, education, and earnings outcomes for the 2000 cohort of college applicants in Taiwan. The micro-data are merged using national IDs. Table 1 provides the descriptive statistics. The sample of college admitted students appears to be evenly balanced in gender and entering college at the expected age of around 18. On average, their parents had less than a high school diploma when the student was born, which is consistent with the historical expansion of public secondary education in Taiwan. Although there has been worry that the market for college education has been expanding too quickly and thus lowering starting salaries, a cursory comparison of family income in 2000 and the student's first year earnings after graduation shows that college graduates' earnings are still high, compared to their parents. The average first-year earnings for newly graduated students are only slightly lower than the average of their parents' combined income at the time the student took the college entrance exam. Thus, a college degree still appears to have value in Taiwan.[2]

1.2.1 Admissions data

The main data set is the college exam and admissions data from the 2000 application season (for students admitted to college for the 2000-2001 academic year), which contains all MOE records of exams taken for early and regular admissions. These detailed records include scores for each of the subject tests t that the applicant i took for either early or regular admissions. The February exams are made up of 5 subject tests, each with a maximum score of 15 so that the highest possible score is 75. The July exams are made up of either 5 or 6 100-point subject-specific tests. The specific set of subject-tests a student is required

[2]This is a future project.

to take is predetermined by the MOE for each educational track a student could pursue.[3,4] Thus, depending on the educational track a student pursues, students could earn a maximum of either 500 or 600 points on their set of required tests. To capture in a common measure the performance of students in same college but different educational tracks, I use the percentage of total possible scores: this is calculated by dividing the individual scores by the total possible (either 500 or 600, depending on what was required by the MOE) and then multiplying by 100. From this, peer quality (denoted as $\overline{P_i}$ in the formula below) can be calculated by averaging the percentages at either college- or department-levels. The value of n in the formula's denominator changes depending on whether the college or department level is used.

$$\overline{P_i} = \frac{\sum_{i=1}^{n} \left[\left(\frac{actual\, score_i}{total\, possible\, score_i} \right) * 100 \right]}{n}$$

In the case of early admissions, the data include the department ID and name (including the college's name) where the student applied to, as well as whether the applicant was offered a seat and accepted it. In 2000, 57,820 students took the February exams for early admissions. This is approximately 47% of all students who took part in the 2000 application season. Of those who applied, 7039 or 13.5% were offered early admissions, which comprises about 5% of all who competed in either early or regular admissions in 2000. Only 438 (5.9% of those offered early admissions) rejected their offers for early admissions.

For those remaining who were not enrolled via early admissions, the regular admissions data includes the department ID and name (including the college's name) where the applicant was admitted, if admitted at all. Approximately 111,150 applied to regular admissions, of

[3]See Table A1 in the Appendix for a detailed breakdown of which subject tests are required for each educational track.

[4]Educational tracks can be divided into 8 fields of study, 23 disciplines, or 158 college majors. An academic department within a college teaches one college major.

which 67,898 or 61% were admitted.

1.2.2 Birth certificate data

I am able to match 100% of the individuals in the MOE data to their birth certificate data using the universe of all births in Taiwan for the birth cohorts 1978-1983. The birth data are collected by the Ministry of the Interior Affairs. Of the data collected on the birth certificate, I use the following variables: year and month of birth, child's sex, and parental years of schooling at time of birth.

1.2.3 Earnings data

The administrative data I use on earnings is for the years 2004 to 2005; 2004 is the first year graduates would have been able to find a full-time job. The data is administered by Taiwan's Bureau of Labor Insurance and collected by the Ministry of Labor Affairs. It is an employer-employee matched data set, which includes the employee's national ID and monthly earnings for that year.

As the name indicates, the Labor Insurance data (hereafter "the Labor data") is a record of healthcare insurance provided through employers. In 1995 Taiwan established universal healthcare coverage. All employed individuals are given access to this universal healthcare through their employer, and the monthly premium is split between the employer and the employee. The employee's share is based on a percentage of his/her monthly income. Thus, all employers are required by law to submit monthly earnings records for their employees in order for the monthly premium to be removed from their monthly earnings (Chou et al., 2003).

The data contain the complete universe of all private-sector workers in Taiwan, approximately 82% of the working population. Of those working adults not in the Labor data, about

10% are public-sector employees, 3% are farmers/fishermen, and 5% are self-employed. The former three maintain a different form of public insurance that pre-dated the 1995 universal healthcare. All unemployed (unemployment rate is about 3-4% at this time) are not in this data because their unemployment benefits cover their healthcare provision (Chou et al., 2003).

A significant limitation of the Labor data to which I have access is that only about 33% of those in my sample show up in the earnings data for 2004 and 2005. This could be due to multiple reasons consistent with the average state of the labor market at this time. The main reason for the lower match rate may be due to gender-specific binding obligations. Males who graduated college in 2004 would have had to perform their 18-month national military service at this time — unless they were, for example, deemed physically unfit, were an only child needing to care for elderly parents, or had an older brother serving at the same time.[5] Because males who graduate in June 2004 would just be completing their 18-month military service in December 2005, the earliest we could expect them to employed would be January 2006. Females may decide to marry following college graduation; in which case, childbearing may immediately follow marriage since that is a common expectation in Asia. Because the military service obligation is more binding than cultural expectations, this would somewhat explain that there are twice as many females than males that could be matched with their wages. Thus, a desirable solution would be to get the Labor data for 2006; unfortunately, access to this data is not currently possible.

As mentioned above, a second reason for the lower match rate is that about 10% of all working adults hold a position in the government, which could include bureaucrats, public school teachers, or public service workers, to name a few. These public servants would not

[5]Reasons for not performing the service immediately (or at all) following graduation from college would be continued education at the post-graduate level, being deemed physically unfit for various reasons, being the only child and thus would leave the parents without a child to assist them, having another family member currently servicing in the military, or fulfilling a bond with a government research position.

appear in the Labor data. Other reasons explaining the lower match rate are that graduates went on to graduate school (though this is even more restricted access than college in Taiwan) or they are unemployed (the unemployment rate is around 4% for college and university graduates at this time) (Yang et al., 2010).

To determine whether this data limitation signals some element of selection, I check for balance in the probability of having wages on either side of the discontinuity. If we were to find a statistically significant (positive or negative) effect of having scored above the admissions threshold, then this is information that could influence our interpretation of wage differentials later on. If the effect were positive as well as statistically significant, then we would be concerned that any wage impact we observe would really be due to selection into finding a job in the private sector. However, since public-sector jobs tend to have better benefits and higher wages and are typically more competitive than the average private-sector job, we may be less worried about selection into the private sector and more worried about selection out of it into the public sector: a concern we might have if we were to find the probability of having wage observations in the Labor data to be negative and significant at the discontinuity. Panel A of Figure 4 shows, however, that there is balance in probability of showing up in the Labor data on either side of the discontinuity. Thus, selection into employment and sector do not appear to be driving our reduced-form results.

Despite our preference for a 100% match rate, possessing 33% of the complete universe of college graduates from the 2000 cohort is still a considerable advantage over most analyses that rely on census or survey data. These studies rely on selected subsamples of the whole population; thus, if we were to focus in those data solely on the graduates from the 2000 cohort, the number of observations that would be available are considerably fewer. Therefore, I take the 33% match rate and the evidence for no selection in or out of the Labor data as positive contributions. I do run a separate regression for women only to test for sensitivity. This substantially reduces the sample size in these regressions, so the results will be noisier.

Because the Labor data is generated through the bureaucratic structures of the universal health-care system, it retains two key features of that system: it is measured somewhat coarsely and at monthly intervals. The coarseness of the Labor data can be seen in Figure 2. The upper limit in our records is 42,000 NTD, which is sufficiently high for the earnings of most new graduates. Any wage above 42,000 were top-coded at the upper limit by the Bureau of Labor Insurance. Similarly, the lower limit is 11,000 NTD, which again is well below the average earnings. Any wage below 11,000 were bottom-coded as the lower limit. Lastly, monthly earnings are listed in lumpy intervals, not in single-dollar units. Though this is not a continuous measure, it is nevertheless measured at fairly consistent and reasonable intervals of around 600 NTD.[6] These characteristics of the Labor data — the truncated distribution and the lumpy intervals — are artifacts of the fee schedule for healthcare with a floor, cap, and menu-pricing scheme.

The Labor data I have access to is taken from the employee earnings records for the month of December for each year. There are advantages to this feature of the data. The first advantage of having monthly earnings from December is that we can capture earnings soon after graduates complete their education in June. A second advantage is that the data is not self-reported, but rather part of a uniform administrative process. Self-reported earnings, which are common in survey data, have been shown to be less precise measures of actual earnings. A third advantage over monthly earning records in survey data is that we record all observations on the same month of the year, which protects against the noise induced by business-cycle seasonality in monthly earnings. A possible disadvantage, however, is that if a college graduate were between jobs in either December 2004 or December 2005, the data would not capture his/her employment. The only way to remedy this would be to have the monthly earnings for every month. Unfortunately, it is politically impossible to obtain this

[6]The lumpy intervals for wages in the Labor data may contribute to the wider confidence intervals we see later in the wage regressions.

extent of data.

First-year earnings, as opposed to earnings a few years after graduation, is a theoretically appealing measure for this study because it would reflect an employer's reliance on college quality as a signal of students' potential productivity. If employers use college quality as a signal to predict an applicant's expected productivity, then first-year earnings should follow a similar pattern as the variance in the signals of college quality the employers observe(MacLeod and Urquiola, 2011). Another way first-year earnings could be useful in this study is as a proxy for longer-term earnings. Oreopoulos et al. (2012) show that a college graduate's starting salary is predictive of long-term earnings growth; thus, on average, starting off on a lower (higher) rung of the pay scale means that longer term earnings will also be lower (higher).

1.3 Empirical Strategy

The counterfactual I construct embodies the basic regression discontinuity design: I use a student's total score on the February exams as the running or forcing variable (x-axis) in the RD. A student's score is either above or below the cutscore c for admissions into that department. The RD design works in the spirit of a randomized experiment in that those who are just above and just below the admissions threshold are essentially the same in observable characteristics, most obviously in their nearly identical test scores. In many cases where admissions cutoffs are set based on test scores, the random measurement error inherent to the construction of standardized tests provides enough random variation to support the claim that students on either side of the cutoff are identical *in expectation*. Thus, those who are just below the threshold provide a good comparison group for those who are just above, and the differences in outcomes that occur later are much more plausibly argued to be due to, in this example, being admitted to a better college. Because of the interview portion of the early admissions protocol, the RD framework I have constructed is a fuzzy RD design, rather than a sharp one. Thus, I interpret the results as an intention-to-treat outcome, rather than a treatment-on-the-treated.

1.3.1 RD with Single Discontinuity

Consider a hypothetical case in which a department has more applicants than there is capacity. The cutoff c_1 defines the access threshold to that department, d_1. Students scoring above the cutoff c_1 are considered for early admissions into d_1; those not admitted early must apply for regular admissions. This constitutes an RD for one department. Model (1) below represents such a hypothetical case:

$$Y_i = \beta 1(T_i \geq c_d) + b(T_i - c_d) + q1(T_i \geq c_d) * (T_i - c_d) + X_i + \varepsilon_i \tag{1.1}$$

where $1(T_i \geq c_1)$ is an indicator function of whether a student's February exam test score is above the cutoff for consideration at d. The function $b(T_i - c_d)$ is the running variable for the February exam scores recentered around the cutscore for that department, independently flexible above and below cutscore through the effects of the interaction term $1(T_i \geq c_d) * (T_i - c_d)$. Within a narrow enough bandwidth around the cutoff, $b(T_i - c_d)$ should capture the effect of other factors driving admissions at d. Because the February exams are measured in discrete whole numbers, I use robust standard errors and cluster the error term ε_i at the test point level. Dong (2010) shows that clustering robust standard errors on the values of the running variable accounts for specification errors that may occur by having a running variable that is discrete instead of continuous.

Notice that the outcome Y_i in equation (1) is fundamentally a function of the relative distance a student's test score is from the cutoff. The intention-to-treat (ITT) interpretation of the β coefficient represents the reduced-form change in some outcome Y_i (whether academic or labor market) that i experiences from scoring just above the cutscore compared to just below. A set of student characteristics X_i are included, but are not required for the β coefficient to be causally identified. An identifying assumption of the RD design is that students on either side of the cutoff are otherwise similar in their background characteristics. Though it is fundamentally impossible to check whether all student characteristics are balanced, I test for balance in the characteristics I have access to by inspecting whether there is overlap in their confidence intervals on either side of the discontinuity. In Figure 4 Panels B-F we see that the confidence intervals for each of these covariates do indeed overlap, indicating that the students are similar across the cutscore. Table 2 shows these same results in tabular form. The graphical and tabular results agree. Mother's years of schooling does seem have some marginally significant effect (at the 10% level) for those who score just above the cutoff. However, if we focus on the bin-averages (dots) in Panel F of Figure 4, it is evident that those students just to the above and below of the cutoff have mothers with nearly the same education levels. In either case, I show below (in Table 6) results with and without covariates included; the outcomes are essentially same with the without-covariate results being slightly stronger in magnitude and significance.

Conditioning access to a department on the relative distance (above or below) a student's test score is from a cutoff provides the random variation necessary to econometrically circumvent bias in our analysis. In Taiwan the cutscore for the early admissions process is the ex-post realization of factors beyond the control of any one student or admissions committee member in the department. Factors governing the placement of the cutscore are the number of those who apply to that department, what their February exam scores are in relation to each other's, and most critically, the capacity constraint set by the MOE for that department. These are all exogenous to the individual student and the admissions committee members. Another factor that is exogenous to the student is how his/her own ability matches the idiosyncratic preferences for student quality within the admissions committee.

It is the combined effects of these factor that create enough randomness in the cutscore placement to identify a causal effect.[7] In the Taiwan setting, the cutscore is a function of the predetermined and binding capacity constraint set by the MOE. US colleges do not have these strictly set and observed constraints in early admissions cohort sizes. The size of early admissions cohorts in the US can be almost completely elastic to the supply of applicants and is determined solely by the college. Therefore, for admissions at US colleges, a valid cut-score can only be set before admissions begins (Lee and Lemieux, 2010). In Taiwan the capacity constraint is both pre-existent and set by an outside agent, and is what forces a cutscore in the first place. Consequently, I argue that Lee and Lemieux (2010)'s description of a valid cut-score also applies here.

1.3.2 Stacked RD with Multiple Discontinuities

My analysis exploits 689 discontinuities, each one its own valid RD. Instead of treating each discontinuity as an individual natural experiment with its own local average treatment effect (LATE), stacking the RDs allows me to aggregate all the discontinuities into one natural experiment and average all the LATEs into one estimate. Besides simplifying the task of interpreting the results, stacking RDs has the desirable characteristics of increasing statistical power by pooling observations around the discontinuity; it also diminishes the bias from extreme outliers. Model (2) updates Model (1) to accommodate the multiple discontinuities. A fixed effect $\gamma(c)_d$ is included to capture the fact that the cutscore c_d for each department occurs at different points in the distribution of February exams.

[7] "Most colleges have multiple members of an admissions committee [....] Consequently, applicant ratings contain random variation due to differences in raters' opinions and variation in their opinions over time. [....] Lee and Lemieux (2010) demonstrate that such random variation is the sole factor determining which candidates fall just below and above a cut-point. They thereby demonstrate that imprecise control over ratings is sufficient to produce random assignment at the cut-point, which yields a valid RD design, as long as the cut-point is not chosen with knowledge of the candidates' ratings" (Jacob et al., 2012, p. 42).

$$Y_i = \beta 1(T_i \geq c_d) + b(T_i - c_d) + q1(T_i \geq c_d) * (T_i - c_d) + \gamma(c)_d + X_i + \varepsilon_i \qquad (1.2)$$

Recentering takes place at the department level; for each department the exam scores T_i for each student i who applied there are normalized around the cutoff c_d for that department. This aligns the discontinuities for each department at T=0. After recentering, stacking the data for each department causes the density of recentered test scores to peak at the cut-score. (See Figure 3.) This is a mechanical feature of recentering and then stacking. Despite the fact that there will not always be an applicant in each department for each point in the test score distribution, recentering the test scores for each department forces each stack (department) to always have at least one observation at T=0. The density at the cutscore, therefore, artificially increases; it has nothing to do with students being able to manipulate which side of the cutoff they land on. Nevertheless, to allay any doubts, I follow the strict recommendations of Barreca et al. (2011) and run a 'donut-RD' by dropping observations at T=0.[8]

1.4 Main Results

Following Imbens and Lemieux (2008), my main results are cross-validated across a wide range of bandwidths at 3-point intervals. The leftmost column in each of these tables has results using a bandwidth of only 1 point on either side of the discontinuity, while the rightmost column uses a bandwidth of 19 points on either side of the discontinuity. For the bulk of my analysis, discussion, and graphs, I settle on a bandwidth of 10 points on either side of the cutoff as providing suitable power, considering that my sample with wages is only 33% of the full. Identifying the bandwidth that minimizes the mean squared error (a

[8]Running the regressions with or without the observations at the cut-score makes little difference to the results.

measure of model fit) is currently a common practice in regression discontinuity design. The Imbens and Kalyanaraman (2009) (hereafter IK) optimal bandwidth estimator[9] is a tool for identifying this specific bandwidth. In this case, the IK optimal bandwidth is just over 4 points. Thus, estimates using a bandwidth of 4 are included (second column from the left) in my cross-validation. Though 58,156 students applied for early admissions in 2000, only the 36,483 students who were admitted to college at all (via early admissions or regular admissions) will be included in the analysis, since my analysis focuses on the returns to college quality, not the returns to a college education in general.

1.4.1 First-Stage Outcomes

Table 3 provides the first-stage impacts on peer quality at the college level in Panel A and department level in Panel B. Since each cutoff is a discontinuity separating access into both a college and a department, the first-stage outcome can be measured for both the college as a whole as well as the department within a college. I start with the peer quality at the college and then consider the variation in peer quality at the specific department level.

Using a bandwidth of 10 points on either side, Panel A, Column (4) shows that students scoring above the cutoff were exposed to peers who, on average, scored about 0.7 percentage points or about 0.05 σ higher than those students who scored below the cutoff. The estimate is statistically significant at the α-level$<$0.01 and serves well as a summary of the range of magnitudes we observe in the cross-validated results: the range includes the lowest magnitude of 0.02 σ in Column (2) and the highest at 0.1 σ in Columns (1) and (7).

The tabular results above match the impacts we see in Figure 5. The 3 panels in Figure 5 give the smoothed fits of average peer quality at the college level around the discontinuity. I provide three different panels as another way to show that these results are robust to

[9] The Stata command for this is rdob.ado, and is available from Imbens' website: `http://www.economics.harvard.edu/faculty/imbens/files/rdob.ado`.

different specifications. For all three panels the bandwidth is 10 points on either side of the discontinuity. Panel A (leftmost) uses the unconditional outcomes. Panel B (center) uses residuals from a regression on a linear trend in the February test scores, including only cutoff fixed effects — that is, no covariates are included. Panel C (rightmost) uses residuals from a regression on a linear trend in the test scores, including covariates and cutoff fixed effects. The dots in each panel represent the within-bin averages for each test point for that specific outcome and specification. Panel C corresponds to the specification shown in Table 3, Column (4). On average, students to the right of the discontinuity were in colleges that had peers scoring about 0.05 σ higher on their exams.

In my analysis, I take into account both college-level and department-level peer quality. Because the quality of the whole institution is possibly one of the most salient factors students account for in their admissions decisions and because college-level peer quality is more likely to be consistent over time, it serves as a good proxy for other relevant institutional characteristics that may attract students. However, it is also common to hear anecdotes of how colleges specialize and are known for their work in a particular discipline, which is a department specific quality outcome. Thus, to take into account these different levels of peer quality, I re-run the same regressions, this time using the department-level peer quality. Table 3, Panel B provides the results. Across the bandwidths in the cross-validation, the results are smaller in size and significance, especially at the more narrow bandwidths. Using preferred bandwidth of 10, the results are statistically significance at the α-level$<$0.05. These coefficients correspond well with what we see in Figure 6, with Panel C corresponding to results in column (11) of Table 3, Panel B. Comparing just the coefficients at the 10-point bandwidth, the impact of scoring above the cutoff is weaker at the department level than the college — only about 60% the size of their corresponding college-level coefficients in Table 3, Panel A. Yet, an overall impact of 0.04 or 0.05 σ is still present for both the peer quality at the college and department levels.

While Table 3 focuses on academic measures of peer quality at the college and the department levels, Table 4 provides results for another salient measure of college quality: prestige or college rank. In the system of higher education in Taiwan, as in other non-Western countries (Rubinstein and Sekhri, 2011), national universities are more prestigious than private institutions. They also offer lower tuition because of the government subsidies. Thus, national universities tend to attract the most academically competitive students. In terms of my research design, the confluence of these benefits make it impossible to disentangle the motivations for wanting to attend a national college. However, I am able to measure the differential impact of scoring above the cutoff on the probability of being admitted to a national college. Further, among the prestigious national colleges, there are an elite 6 (top-6) for which I also measure the admissions impact of scoring above the cutoff.

To measure the prestige (or institutional) effect, I use a simple 0/1 indicator for admissions into a national or top-6 institution as an outcome. If employers view the label "National" in the college's name, or even the specific name of a top-6 institution (like National Taiwan University) as an important signal of quality, then having this label could have a differential effect in setting wages. Table 4 provides the results for both the probability of admissions to a national college (Panel A) and admissions into a top-6 institution (Panel B). Across the bandwidths, the differential probability of being admitted to a national college increases by about 4.3 percentage points, which is is equivalent to a 0.1 σ improvement. When compared to the population mean of 20%, the impact again seems considerable. These results are highly statistically significant and the discontinuity is visually stark in Figure 7. (Panel C in Figure 7 corresponds the results in Column (2) of Table 4.) Table 4, Panel B shows the impact on attending an even more prestigious rank of college, the elite top-6 national college. The coefficients across the bandwidths are very similar to those in Panel A. In Column (11), the impact of scoring above the threshold is a 3.3 percentage point or about a 0.12 σ increase in admissions to a top-6 institution. Figure 8 Panel C provides the corresponding visual evidence.

The positive impact on the rank/prestige outcomes suggest that students above the cutoff are, on average, exposed to better quality colleges than those below. National colleges are known for better faculty, more diverse college majors, and better facilities, and as my results show, higher performing peers. These prestige differentials could also be proxying for productivity differences between colleges, as MacLeod and Urquiola (2011) argue.

1.4.2 Reduced-Form Outcomes

Tuition at a national college can be about half of what it is in a private college. Tuition can also vary within colleges at the department level, with science-track college majors costing twice that of social-track ones. Therefore, I look at the differential returns to tuition paid as a relevant return to college quality. In the preceding section, I treat peer quality at the college and department level, as well as probability of admissions into a national or top-6 institution as first-stage results because they all occur during the admissions process.[10] Even though the tuition that students pay is set at the same time as the previous four outcomes are realized, tuition is paid each year for four years and thus affects long-term savings/spending of the families resources over the medium-term. Therefore, I treat it as a reduced-form outcome along with wages.

Table 5, Panel A gives the cross-validated results for the log of yearly tuition paid. In line with the first-stage results we observed above, especially for access to the national institutions, students who scored just above the threshold for access were, on average, paying about 3.1 log-points *less* tuition each year, compared to those who scored below the threshold and had to be admitted via regular admissions. The savings to family resources amounts to nearly a 0.1 σ decrease in yearly tuition, which is the magnitude we would expect, given

[10] I do not have measures of actual attendance and completion of college degree.

the 0.1 σ increase in access to national and top-6 colleges. In real dollar terms, this is about 1100 NTD or about 3 to 5% of a family's combined monthly earnings in 2000.

Moving to the labor market outcomes, Table 5, Panel B shows the reduced-form effects of scoring above the cutoff on first-year monthly earnings. The results are substantive. In Column (4), we find a highly statistically significant estimate of about 3.2 ppts in log wages, which is equivalent to about 0.1 σ increase or about 700 NTD more per month. This discontinuity matches what we see graphically in Figure 10 in Panel C.[11] Notably, the impact of scoring above the discontinuity is relatively stable across the range of bandwidths; despite the reduced sample size, the results are statistically significant at conventional levels. Table 6, Panels A and B show the wage results with and without covariates included. These regressions agree in size, sign, and significance.

The magnitudes of the reduced-form returns compare quite favorably with other similar studies. In the better schools literature, Pop-Eleches and Urquiola (2012) is most comparable for its similar research design using a stacked RD. They find that attending a better high school is associated with a 0.02 to 0.09 improvement in college entrance exam scores 3 years later, depending on whether the impact is measured with administrative or survey data. Of the college quality papers, Saavedra (2008) is most comparable for measuring the impacts at multiple discontinuities. He finds that attending the most selective institution in Colombia is associated with 0.25 σ increase in performance on a college exit exam and about 35% increase in first year earnings. Regarding the earnings benefit, Saavedra points out that these are unconditional regressions in that they include those with and without employment. For those without employment, a $0 amount is imputed; thus, the 35% increase is capturing both the impact on employment and the impact on wages. Once he conditions his regression on being

[11]The confidence intervals in this graph appear to be wider than the standard errors in Table 4; this is because the Stata smoothing function `lpolyci` does not account for the effect of clustering the robust standard errors in the regressions.

employed, there appears to be no significant difference in log-wages for those admitted to the selective institution versus not. In my analysis, all wage regressions are conditioned on being employed. Because of the different reasons for students not showing up in the Labor data (discussed in detail in Section 2.3), I cannot impute zero earnings.

As mentioned before, the size of the first-stage results in peer quality seem to be small in their own right, but if we were to attribute all of the earnings gains observed in the reduced form, then this would indicate that even a small change in average peer quality has a substantial impact. MacLeod and Urquiola (2011, p.14) argue that if peer effects were the main mechanism through which college quality had an effect on earnings, then the advantage should be a permanent impact, which is not what we see in Dale and Krueger (2002, 2011), where the earnings benefits fade over time. Unfortunately, I do not have access to wages beyond the first year, so I am unable to test this myself.

1.5 Heterogeneity

In Section 4, I provide evidence that scoring above the admissions cutoff provides students with better peers and prestige (or institutional quality), and that these in turn correspond to increased monetary benefits in the form of lower tuition and higher wages. With these facts established, an important consideration now is whether these returns are heterogeneous. We might expect that the returns to college quality may differ by gender or non-linear in levels of peer quality, types of college majors, or prestige of college. Below I test for these.

1.5.1 Differences in Returns by Gender

To test for differences in returns by gender, I re-run the original specifications separately for males and females. I report the reduced form results in Table 7. Across the bandwidths, the results for females follow the main results in size, significance, and direction. Results toward the middle bandwidths tend to dip off slightly in size and more so significance, but a generally in robust. Across the bandwidths, the results for males tend to be small and weak at the very narrow bandwitdths, which most likely is do to the very small sample size. But once the bandwidth exceeds 7 points, the results are substantive and significant at conventional levels and consistently at around 0.055 log points. Comparing between males and females, it seems that most bandwidths above 7 points, the coefficient on males is twice that of females. Males earning more than females is consistent with the general literature. It is consistent with the traditional aspects of Asian cultures which encourage males toward more prestigious fields (business and sciences) and high-risk, high-reward career, whereas females are encouraged into less prestigious or competitive fields and more stable careers (education or support services) (Buser et al., 2012; Niederle and Vesterlund, 2007).

1.5.2 Non-linearities in Peer Quality: Top- versus Bottom-Half of Colleges

These regressions are constructed by splitting the original sample into 2 subgroups, depending on whether the student applied to a department that was in the top-half or bottom-half of the peer quality distribution within that college major. Status as top half and bottom half is constructed by ranking all departments within each college major by the departmental average peer quality $\overline{P}$, and dividing it evenly above and below the middle. In cases where there is an odd number of departments in the college major, the middle department is given to the top half.

Conspicuous throughout this section is the extent to which students experience consistently better outcomes for scoring above the cutoff if they applied to a department in the top half of the peer quality distribution. Applying to a bottom-half academic department provides no differential quality benefit. For the figures in this section, the two panels give the smoothed fits of specific outcomes around the cutoff. Panel A (left) provides the results for those who applied to the bottom half in peer quality. Panel B (right) provides the results for those who applied to the top half. Table 8 provides all the results.

Figure 11 shows the outcomes for peer quality at the college level. Scoring above the cutoff improves students' peer quality by 1.19 percentage points if they applied to the top-half of departments (Panel B), but does nothing for those applying to bottom-half departments (Panel A). This is exactly what we see in Table 8 Columns (1) and (2). Similarly, Table 8 Columns (3) and (4) and Figure 12, Panels A and B show that scoring above the cutoff for the top-half departments exposes students to differentially more able peers at the department level as well by about 1.07 percentage points. The prestige effects, shown in Figures 13 and 14, are also noticeably non-linear in peer quality. Students applying to top-half departments within their major experience a 7.8 percentage point (0.078 proportional change) in the probability of being admitted to a national-ranked college (Column (6)) and a 5.1 percentage point increase in the probability of being admitted to a top-6 elite national college (Column (8)). Lastly, the monetary benefits to applying to and scoring above the cutoff at a top-half department line up with what we saw above. Figures 15 and 16 show that students pay less tuition by about 5.5 log-points per year (Column (10)) and earn about 4.7 log-points higher wages per month (Column (12)). Those at the bottom-half experienced no monetary benefits for scoring above the cutoff.

Across the six outcomes (peer quality at the college- and department-levels, national and top-6 college rank, and log-tuition and log-wage returns), the effects of scoring above the cutoff for top-half departments have been about 60% larger than when the sample was pooled together. While there appears to be little differential benefit to scoring above the cutoff for bottom-half departments, there may be unobserved benefits that motivated students' to apply these departments. For example, if college itself has consumption value, then improving the chances of matriculation, even at lower quality departments, could be an advantage (Oreopoulos and Salvanes, 2011).

1.5.3 Non-linearities in College Rank and Major

To test the effects of college major and college rank together, a desirable research design would be to randomly assign students to one of four academic departments, based on the four possible pairings of private versus national college, and social versus science tracks. Then a simple comparison of the average earnings of students in each group would provide the causal estimates of attending a private versus national college, holding track constant, or studying a social- versus science-track college major, holding college rank constant. Given that this is not socially possible, we can use the fact that students apply to both a college and academic department at the same time. Since each academic department within a college is associated with only one of the two educational tracks, and each college is either private or national, we can use this to inspect the heterogeneity in the first-stage and reduced-form outcomes.[12] To do this I split the sample into four subgroups, corresponding to the four possible pairings.

In this section, each figure will have four panels. The four panels will correspond to the following pattern: Panel A (upper-left) provides the results for those who applied to the social track of a private college. Panel B (upper-right) provides the results for those who applied to the social track of a national college. Panel C (lower-left) provides the results for those who applied to the science track of a private college. Panel D (lower-right) provides

[12]Another alternative approach would be to treat everyone within the sample college major as if they all applied to all the colleges that offered that major and then use the cutoff to the lowest ranked national college as the de facto cutoff for access to any national college within that college major. Unfortunately, the design of the early application process does not lend itself to this kind of analysis because students cannot apply early to all the institutions within that major — only one. Therefore, it would be beyond the constraints of the system to construct this counterfactual, and the interpretation of any coefficient derived from such an exercise would have limited meaningfulness, if any. Such a counterfactual construction would be quite appropriate, though, for the regular admissions process in Taiwan. Indeed, this is the research design that I use in my analysis of the regular admissions process (Simpson 2012b).

the results for those who applied to the science track of a national college. All four panels share a common range for their y-axes. Further, the tabular results are shown in Table 9.

Figure 17 shows that students stay in the same educational track no matter if they scored above the cutoff or below. Across the four subgroup options, the probability that a student studies science is completely balanced on either side of the cutoff. This is a useful finding for our study. Because educational track is one of the two dimensions that we have split the sample on, the absence of change in educational track above and below the cutoff means that we can remove this dimension from our accounting when we look at other relevant outcomes.

However, when we look at the outcomes for the other dimension along which the sample is split, we find that students do change between private and national at the discontinuity (Figure 18). Panel A is the exception here. Students applying to social-track majors in private colleges do not appear to be differentially placed into a national versus private college at the discontinuity. For those applying to national colleges (Panels B and D), the pattern observed in Section 4.1 is unchanged. Students with scores above the cutoff are more likely to get into a national college, compared to those who scored below. Students applying to the science track had an even larger differential advantage in getting into a national college than their counterparts in the social track. The discontinuity for national-science is about twice as large as the one for national-social.

An unexpected finding is that students who applied to private-science (Panel C) were differentially more likely to get into a national college if they were below the cutoff. This is the first indication that the national versus private dimension is not completely synonymous with high-quality versus low-quality, though the correspondence is strong. Three of the top-6 private colleges are medical universities, which could explain why even those below the cutoff are still admitted into a science-track major in a national college. Panel C throughout will provide possible insights into the relative importance of college major over college prestige/type in Taiwan.

Figure 19 provides the results for average peer quality at the college level by subgroup. Panel A is again balanced: Students applying to private-social appear to have little difference in any of their outcomes regardless whether they were above or below the cutoff. The fact that there is little differential influence at the discontinuity for most of the outcomes does not necessarily mean that private-social has the lowest returns in wages, compared to the other 3 panels. However, other facts would suggest this. One main reason for the lower returns may be that private-social has the largest share of overall enrollment; it is between 2 to 3 times larger that either national-social or national-science and about 40% larger than private-science. Being less selective would suggest that the differential advantage of being admitted there is limited. For private-science (Panel C), there appears to be no differential impact on college peer quality at the margin for those applying to this subgroup. Therefore, most of peer quality results we saw in Section 4.1 can be attributed to those applying to national colleges. Since cohort size in national institutions is smaller than at private colleges, the reduced tuition itself could create enough incentive for high-ability students to compete for a seat at a national college. The differential impact on peer quality seems to be largest for the science track at national college, relative to the social track.

For reduced-form results, we see in Figure 20 that the differential impacts on log of yearly tuition follow exactly what we would expect. Students getting into private-science (Panel C) paid more for being admitted at an institution that did not offer subsidies. Private-science applicants above the cutscore paid higher tuition relative to those below, possibly both because those schools were more costly, and because students below the discontinuity were more likely to eventually be enrolled into a national-science program, which enjoys the benefits of subsidized tuition. The differential in log-tuition is most likely due to this differential probability of being admitted into a national college and not the higher cost of these private-science programs. I argue this because in Panel A students applying to private-social programs but fall below the cutoff are no more likely to be enrolled into a national college, and thus we see that they do not experience a change in tuition at the discontinuity.

At both national-social (Panel B) and national-science (Panel D) programs, the log-tuition that a student paid was differentially lower for those above the cutoff score. Though the differential is larger in log-points for Panel D, the difference between relative and absolute returns are important because the bases for each are quite different. On average, the tuition for a science-track college major can be about 50 to 100% more than that for a social-track one in the same college. Thus, in absolute dollar amounts, the science-track students are receiving more money by being admitted to national college, but in relative terms the subsidy may be larger for the social-track students because their program costs less in the first place.

Lastly, Figure 21 shows the differential impact on log wages. If our prior assumption were that national college students will outperform private, the upper quadrants upset this expectation. The upper two quadrants for social-track college majors appear to have little differential benefits to wages from scoring above or below the cutoff. Because of the consistent balance across outcomes for the private-social subgroup, we are less surprised about this finding for those applying to private-social programs (Panel A); but for national-social (Panel B), this lack of wage benefit is inspite of the fact that students applying to national-social were differentially more likely to get into national college and to experience peers of higher ability. From this we may infer that employers value less the skills that are associated with social-track college majors, in general.

The converse is generally true, though, for those who applied to national-science programs. Panel D for each outcome reaffirms our prior assumption that national colleges provide a benefit. In general, Panel D works exactly how one might think college rank and science majors perform, especially together. Students scoring above the cutoff at national-science programs are differentially more likely to be in a national college, experience higher quality peers at the college level, pay lower tuition, and earn higher wages following graduation.

It would appear, however, that the wage differential experienced by national-science students (8.366 log-points) is outdone by the wage differential experienced by private-science students (12.498 log-points). Importantly, these different wage impacts share same log-wage base of 10.548. To pursue more in-depth the reasons for the higher wage differentials in private-science and national-science programs, I re-run the analysis for only private-science and national-science programs on the following outcomes: the probability that the student ends up in a science track, a medical science college major, or a top-6 (national versus private) institution. I also try to tease out the effects of prestige versus college major by noting that several of the top private colleges are medical schools. Thus, for the private-science departments, I also include as an outcome whether the student was admitted into a private medical college or not. Results are shown in Table 7, Panel A for private-science and Panel B for national-science.

The four panels in Figure 22 give the smoothed fits of four separate outcomes for the subgroup of private-science applicants. Panel A (upper-left) provides the results for the probability of being admitted at all into a science-track (versus social-track) college major. This is the same outcome that we saw in Figure 17, Panel C. Students tend to stay within the educational track. Further, when we look at Panel B, we see that students are no more likely to study medicine if they applied to a medical-science academic department. Despite the fact that they are no more likely to study medicine, they nevertheless still are more likely to study in a medical school (Panel C) as well as a top-6 private college (Panel D). Since several of the top private colleges are medical schools, I am unable to distinguish the effect of simply going to a medical school versus an elite private college. However, because the students are no more likely to study a medical-science college major on either side of the discontinuity, we could argue that the differential wage returns to scoring above the discontinuity is more likely a prestige/institutional effect, since they are also no more likely to study science either.

The four panels in Figure 23 give the smoothed fits for four separate outcomes for only the subgroup of national-science applicants. Panels A, C, and D are reproduced here from previous figures for ease of comparison. These three panels show that students applying to national-science are more likely to study science in a national college, and somewhat more likely to do so in a top-6 national college. It appears from Panel B (upper-right) that there is no differential in the probability of being admitted into a medical-science college major by scoring above the cutoff.[13]

Since national-science applicants seem to have first-stage impacts mainly in college peer quality and national rank, with the national rank providing the larger impact in terms of standard deviations, I argue that in general, science-track college majors are sensitive to non-linearities in a way that social-track college majors appear not to be. Whether the sensitivity is to peer quality or to college rank is harder to say. In Section 5.1, I show that science-track majors are sensitive whether they applied to a top-half or a bottom-half academic department within their college major. Those results were not split by college major as well. In this section I show that most of the differential benefits accrue to those in science-track majors. However, within each college major, the private- versus national-college distinction has a strong correspondence to the bottom-half versus top-half distinctions in the peer quality distribution from section 5.1, with the top-half being dominated by national colleges.

To illustrate this correspondence, Figure 24 reproduces the log-wage results of Figure 21, this time splitting the sample into four subgroups by the top- versus bottom-half in peer quality instead of national versus private college rank. (The other dimension the sample is split by is still science- versus social-track college majors.) Similar to Figure 21, the social-track college majors in both Panels A and B are balanced at their discontinuities, whereas the

[13] I do not include the one national medical college in these regressions because all graduates of this program are bonded by the government to work after graduation in a public hospital; therefore, their wages would not show up in the Labor data.

science-track college majors in both Panels C and D evidence a substantial wage differential of about 0.1 log-points for those scoring above the discontinuity relative to those below. Noteworthy is the fact that in both Figures 21 and 24, the wage differentials are larger in Panel C than in Panel D. Since private colleges tend to have their academic departments taking up much of the bottom-half of the peer quality distribution, this would suggest that the larger wage differential supports the claim that excluding the lowest performing peers has greater differential effect than necessarily including the highest performing ones. With only this evidence to substantiate the claim, I treat this interpretation as more suggestive than definitive.

1.6 Conclusion

In this paper I have exploited a novel source of identification for estimating the returns to college quality and college major: Students scoring above the admissions cutoff were more likely to be exposed to better college quality as measured by several proxies. I show that students above the cutoff, on average, experienced 0.04 σ better peers at college, were 0.1 σ more likely to attend a national college, paid about 0.1 σ *less* in tuition, and eventually earned 0.1 σ higher wages in the first year of employment after graduation.

Heterogeneity in returns is quite evident. Students scoring above the cutoff for admissions into national colleges appear to experience much of the differentials in college quality, yet it is only the science-track college majors that experience the wage differentials. Importantly, it is the science-track college majors in both private and national colleges that experience these wage returns, with private colleges experiencing the larger. When the sample is split by whether the academic department (college) falls within the top-half versus the bottom-half of the peer quality distribution for that college major, the results are qualitatively similar to when they are split by national versus private, especially for log-wages. Those in the top-half experience most of the college-quality differentials, but those in the sciences experience all of the wage differentials. Within each college major, there is strong correspondence between an academic department's placement in the top versus bottom half of the peer quality distribution and its type as a national versus a private institution. The similar results between the two methods of splitting the samples (private colleges being mainly in the bottom half of colleges in any given major and national colleges in the top half) would suggest that the labor force is especially attuned to differences in peer quality for the science-track majors, but not for the social-track majors. The larger wage differentials for the bottom-half private colleges may support the claim that wages may be more sensitive to whether your peer groups excludes students with the lowest ability than it necessarily includes ones with the highest ability.

The analysis above remains somewhat agnostic regarding whether the benefits of college quality are actually increasing human capital or only signaling preexistent ability, especially because my research design does not allow me to separate the two. There are arguments on both sides, with most papers though showing evidence for human capital accumulation. Whether or not national colleges helped students learn more or better, they clearly provided a benefit through their subsidized tuition.

This is one of the first papers to estimate the causal returns to college quality using the whole college market of colleges and college students. The evidence above suggests broadly that college quality matters. Peers and college majors are important factors that contribute to the heterogeneity of these general findings.

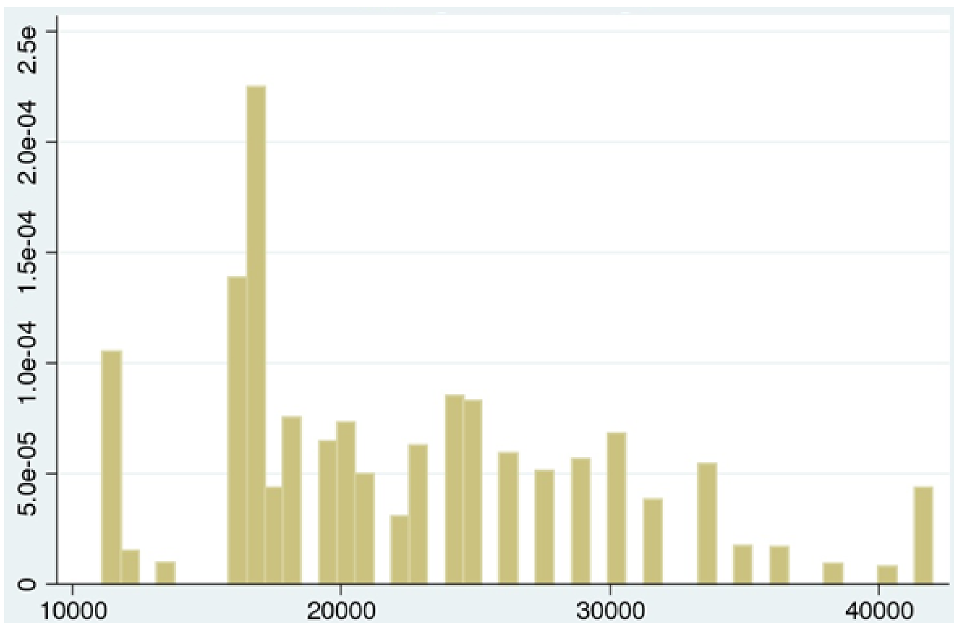
The generalizability of these results may be limited to contexts that are similarly centralized. Though this does exclude countries like the US, it nevertheless includes many other countries that have experienced the largest growth in higher education and, in the case of other East Asian countries, national populations with the world's highest levels of college education. Therefore, this paper offers early and important insights into the causal effects of college quality and college major on earnings.

Figures

Figure 1.1: Timeline of early and regular admissions.

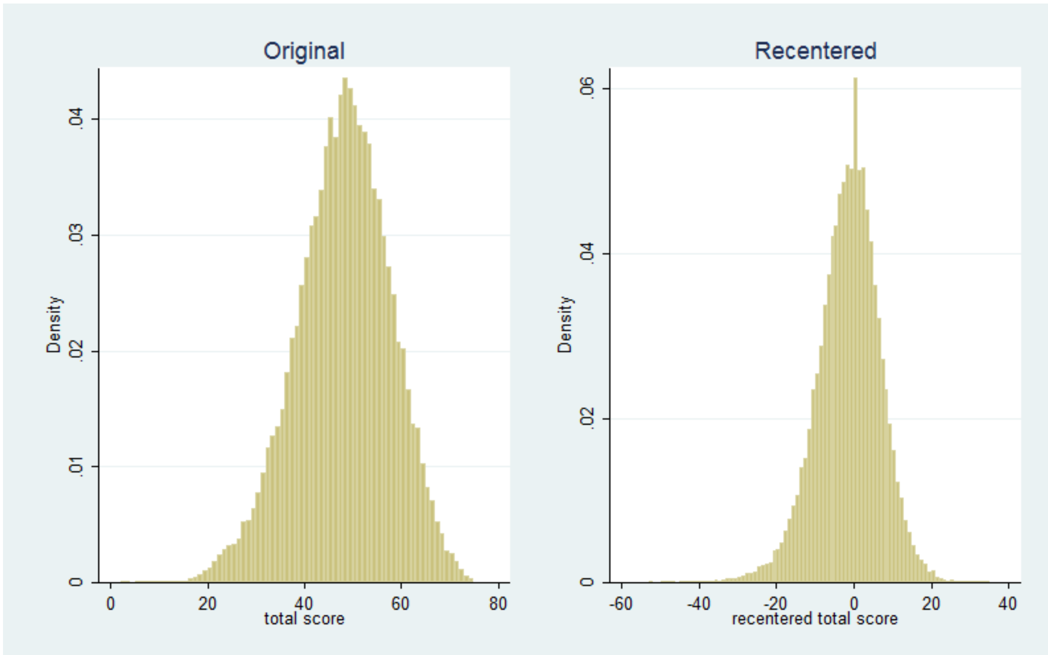
Notes: This graph is an approximation.

Figure 1.2: Histogram of wages



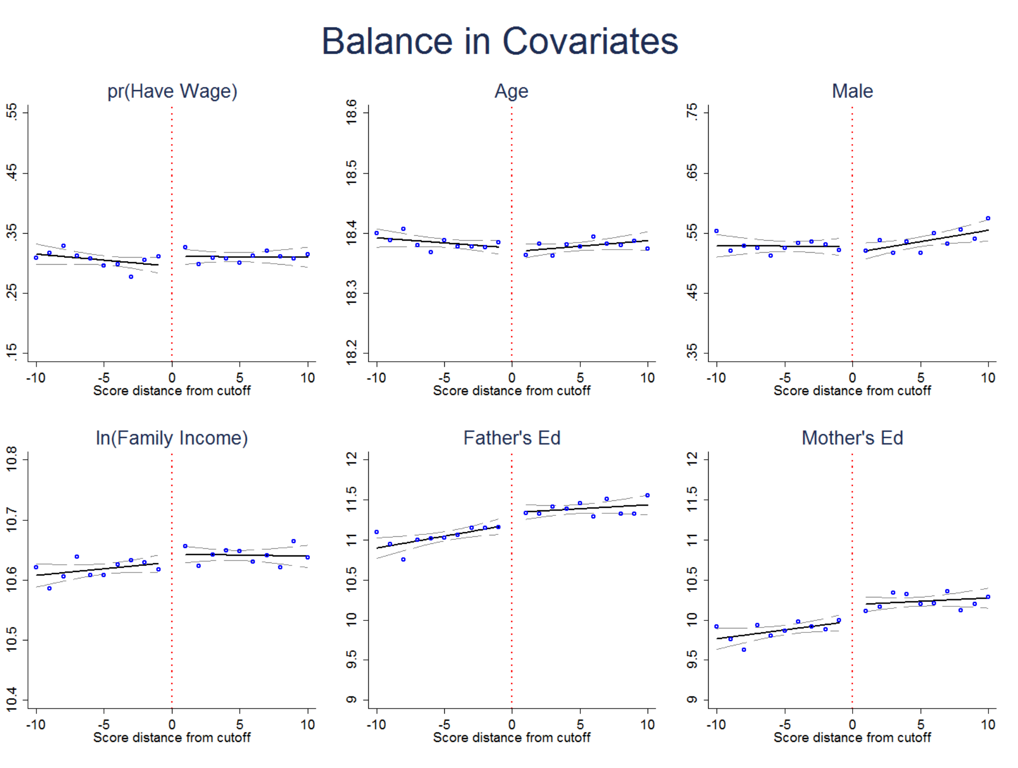
Notes: These represent the first year monthly earnings (in New Taiwan Dollars) for those entering college in 2000. The data were recorded in the month of December for 2004 and 2005. The data are from the Labor Insurance scheme that Taiwan follows in which an individual's premium is a function of his/her monthly salary. Thus, because of the intervals in the menu pricing, the wage records are also lumpy and truncated to a range between 11,000 NTD to 42,000 NTD.

Figure 1.3: Density of Running Variable: Original vs. Recentered



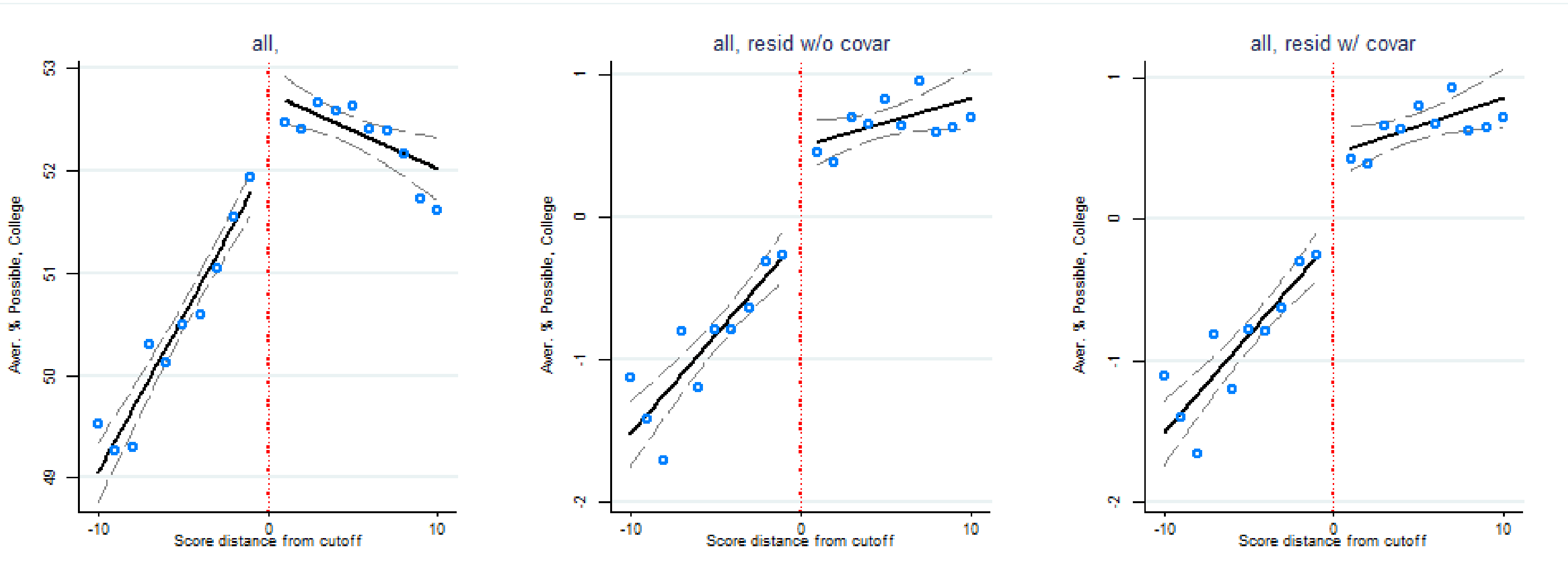
Notes: Process of stacking data mechanically causes heaping at cut-score. I drop observations at cut-score, just in case.

Figure 1.4: Balance in the Covariates



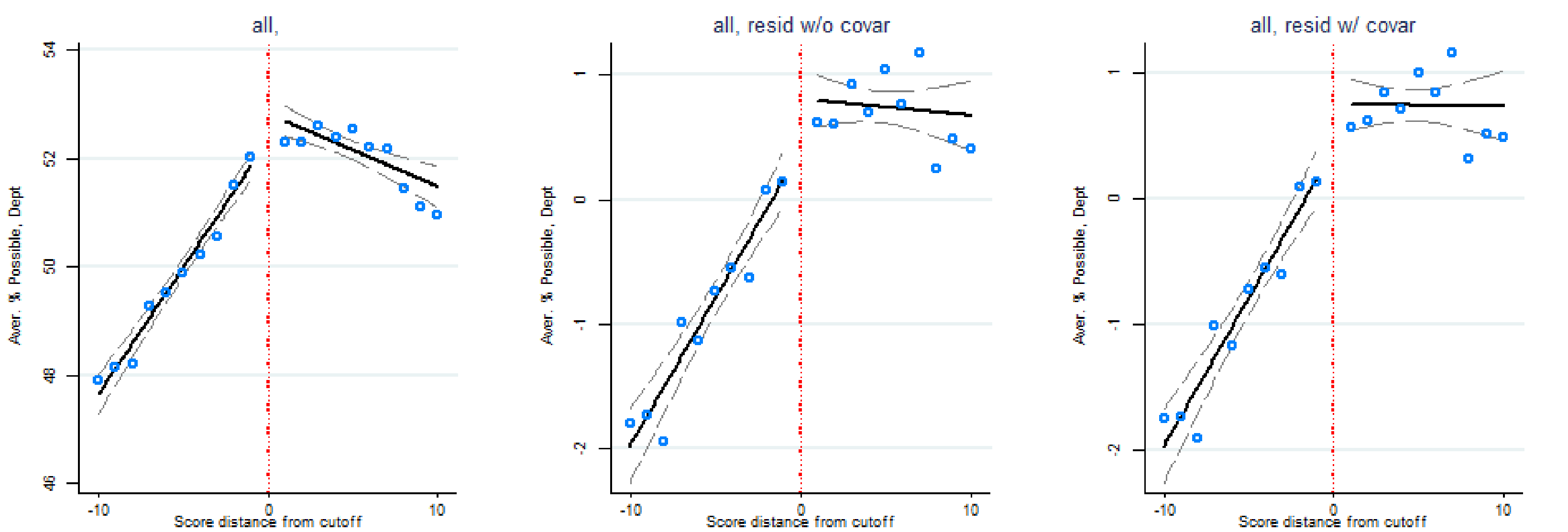
Notes: For each covariate the confidence intervals overlap, suggesting that they are not statistically significantly different from each other at the discontinuity. All panels are linearly smoothed fits on 10 point bandwidths. The dots represent bin averages. Panel A is the probability that an individual is observed with wages in the Labor for the years 2004 or 2005 on either side of the cutoff. Panels B-C are characteristics of the student in 2000 at time of taking the early admissions exam. Panels D-F are characteristics of the parents. Panel D is family income in 2000. Panels E and F are parents' years of schooling at the students time of birth.

Figure 1.5: First-stage: Peer Quality at the College Level



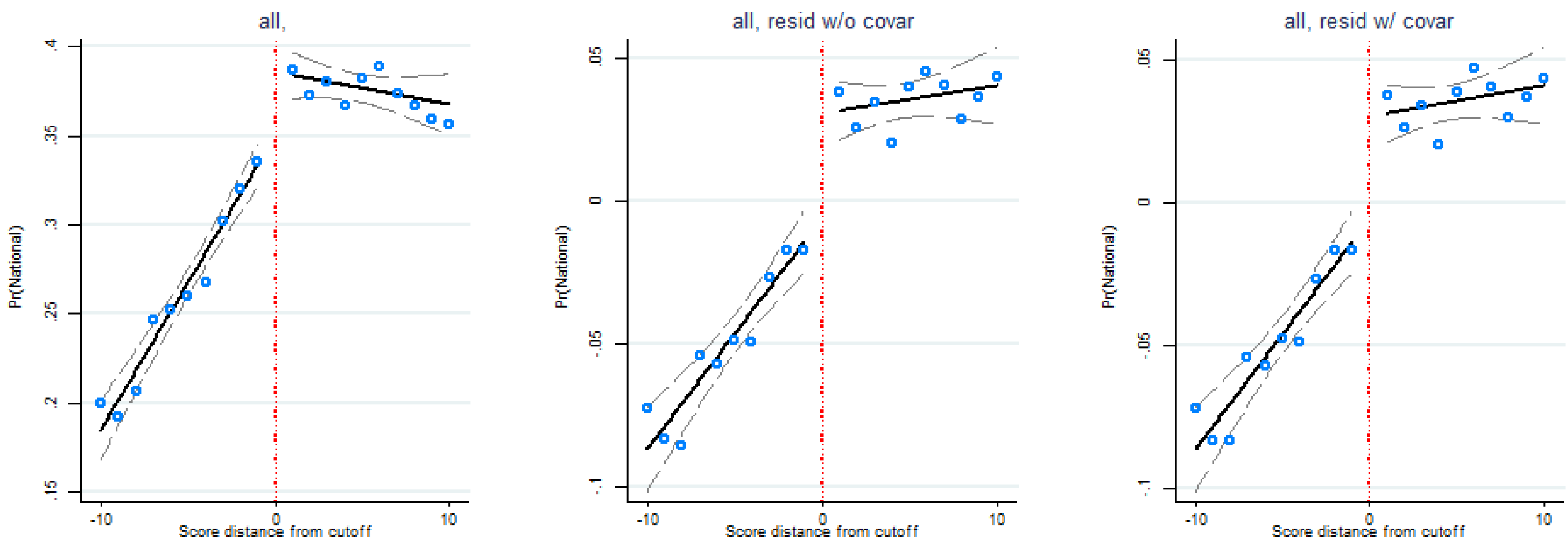
Notes: The three panels above give the smoothed fits of average peer quality at the college level around the discontinuity. The bandwidth is 10 points on either side of the discontinuity. These figures display residuals are taken from a regression of the outcome on a linear trend in the test scores, including covariates and cut-off fixed effects. Only observations for that subgroup are used in the regression. Panel A (leftmost) uses the unconditional outcomes. Panel B (center) uses residuals from a regression on a linear trend in the test scores, including only cutoff fixed effects (i.e., no covariates). Panel C (rightmost) uses residuals from a regression on a linear trend in the test scores, including covariates and cutoff fixed effects. The dots in each panel represent the within-bin averages for each test point for that specific outcome and specification. Panel C corresponds what is shown in Table 3 column (4).

Figure 1.6: First-stage: Peer Quality at the Academic Department Level



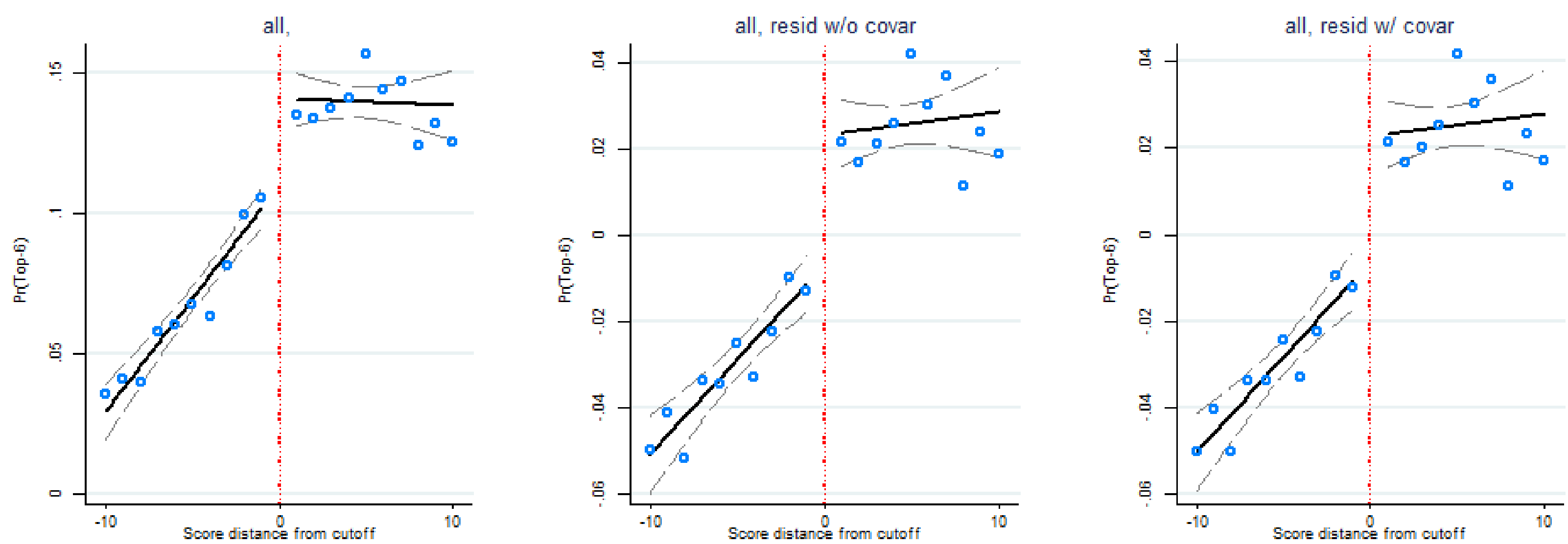
Notes: The three panels above give the smoothed fits of average peer quality at the academic department level around the discontinuity. The bandwidth is 10 points on either side of the discontinuity. These figures display residuals are taken from a regression of the outcome on a linear trend in the test scores, including covariates and cut-off fixed effects. Only observations for that subgroup are used in the regression. Panel A (leftmost) uses the unconditional outcomes. Panel B (center) uses residuals from a regression on a linear trend in the test scores, including only cutoff fixed effects (i.e., no covariates). Panel C (rightmost) uses residuals from a regression on a linear trend in the test scores, including covariates and cutoff fixed effects. The dots in each panel represent the within-bin averages for each test point for that specific outcome and specification. Panel C corresponds what is shown in Table 3 column (11).

Figure 1.7: First-stage: Probability of being admitted into a national (public) college.



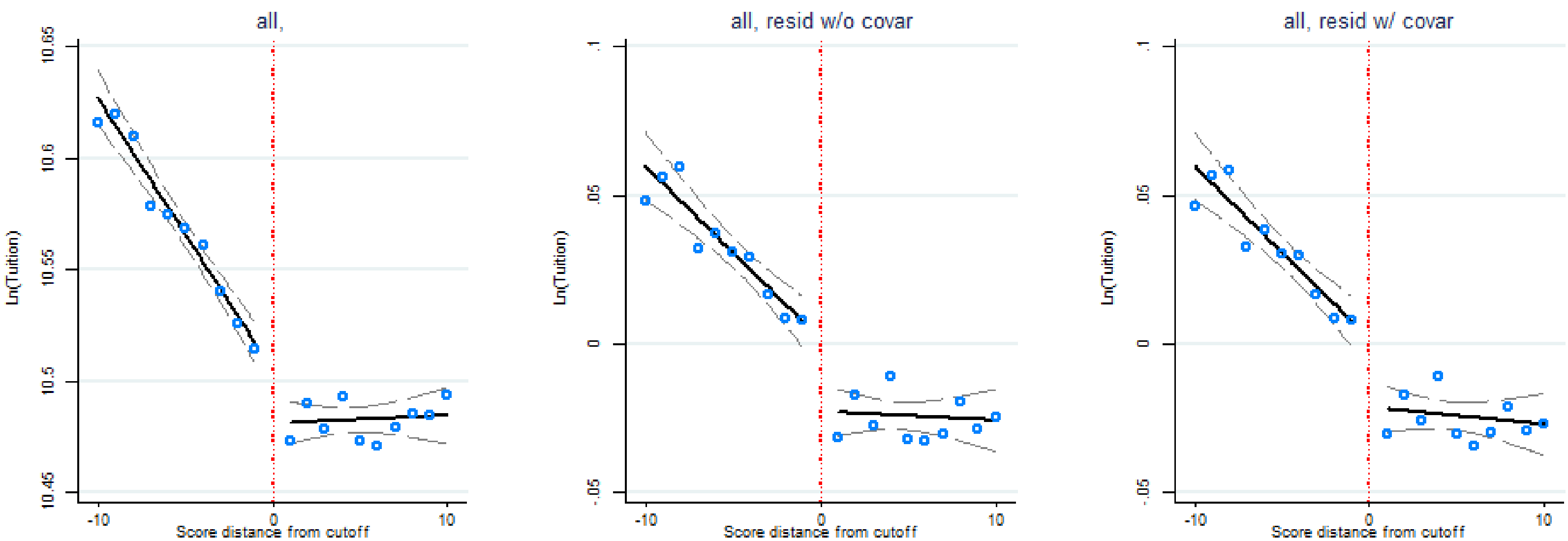
Notes: The three panels above give the smoothed fits of the probability of being admitted to a national college around the discontinuity. The bandwidth is 10 points on either side of the discontinuity. These figures display residuals are taken from a regression of the outcome on a linear trend in the test scores, including covariates and cut-off fixed effects. Only observations for that subgroup are used in the regression. Panel A (leftmost) uses the unconditional outcomes. Panel B (center) uses residuals from a regression on a linear trend in the test scores, including only cutoff fixed effects (i.e., no covariates). Panel C (rightmost) uses residuals from a regression on a linear trend in the test scores, including covariates and cutoff fixed effects. The dots in each panel represent the within-bin averages for each test point for that specific outcome and specification. Panel C corresponds what is shown in Table 4 column (4).

Figure 1.8: First-stage: Probability of being admitted into a top-6 (elite public) college.



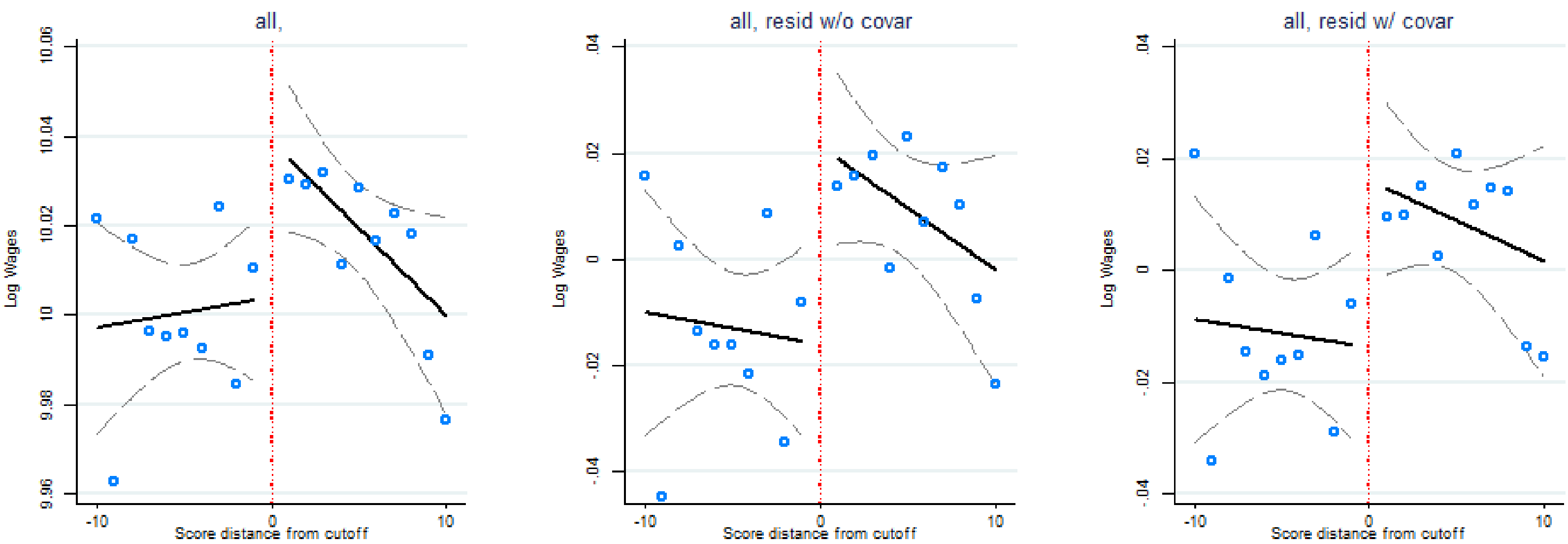
Notes: The three panels above give the smoothed fits of the probability of being admitted to a top-6 college around the discontinuity. The bandwidth is 10 points on either side of the discontinuity. These figures display residuals are taken from a regression of the outcome on a linear trend in the test scores, including covariates and cut-off fixed effects. Only observations for that subgroup are used in the regression. Panel A (leftmost) uses the unconditional outcomes. Panel B (center) uses residuals from a regression on a linear trend in the test scores, including only cutoff fixed effects (i.e., no covariates). Panel C (rightmost) uses residuals from a regression on a linear trend in the test scores, including covariates and cutoff fixed effects. The dots in each panel represent the within-bin averages for each test point for that specific outcome and specification. Panel C corresponds what is shown in Table 4 column (11).

Figure 1.9: Reduced Form: Reduced log yearly tuition.



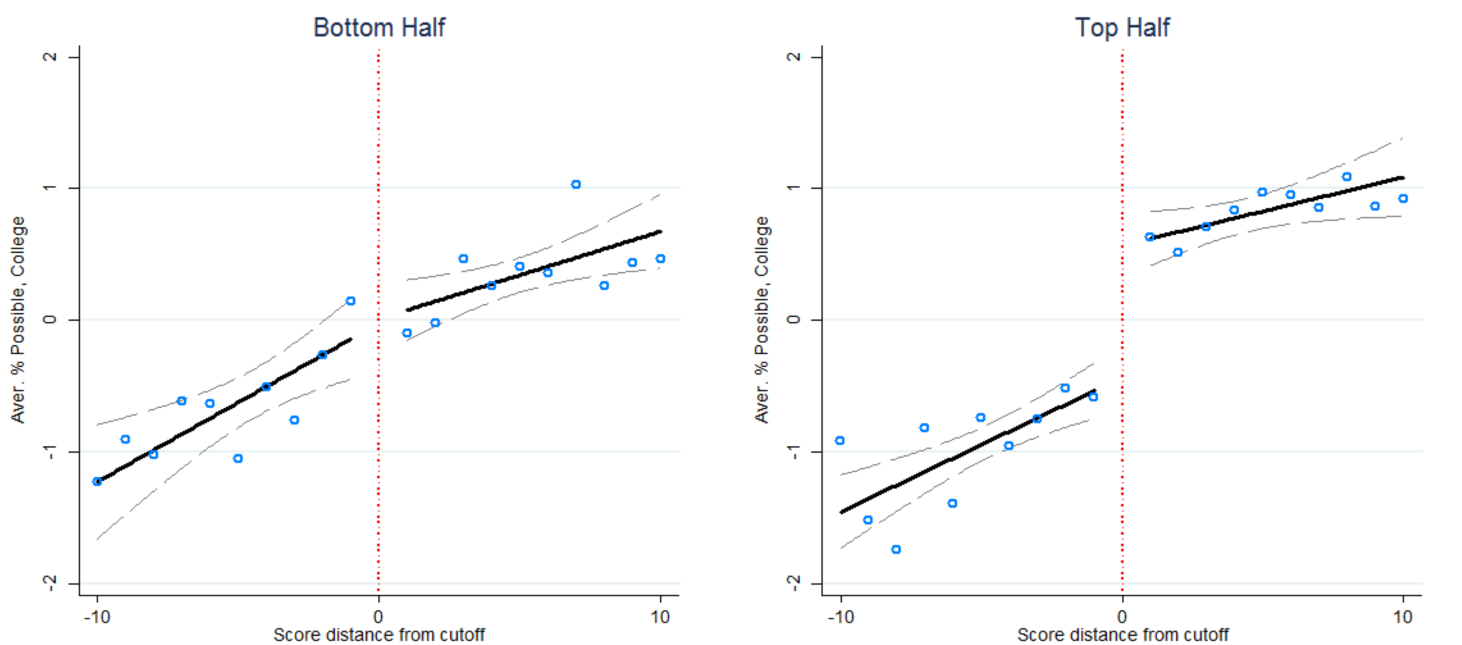
Notes: The three panels above give the smoothed fits of the log of yearly tuition around the discontinuity. The bandwidth is 10 points on either side of the discontinuity. These figures display residuals are taken from a regression of the outcome on a linear trend in the test scores, including covariates and cut-off fixed effects. Only observations for that subgroup are used in the regression. Panel A (leftmost) uses the unconditional outcomes. Panel B (center) uses residuals from a regression on a linear trend in the test scores, including only cutoff fixed effects (i.e., no covariates). Panel C (rightmost) uses residuals from a regression on a linear trend in the test scores, including covariates and cutoff fixed effects. The dots in each panel represent the within-bin averages for each test point for that specific outcome and specification. Panel C corresponds what is shown in Table 5 column (4).

Figure 1.10: Reduced Form: Log wages.



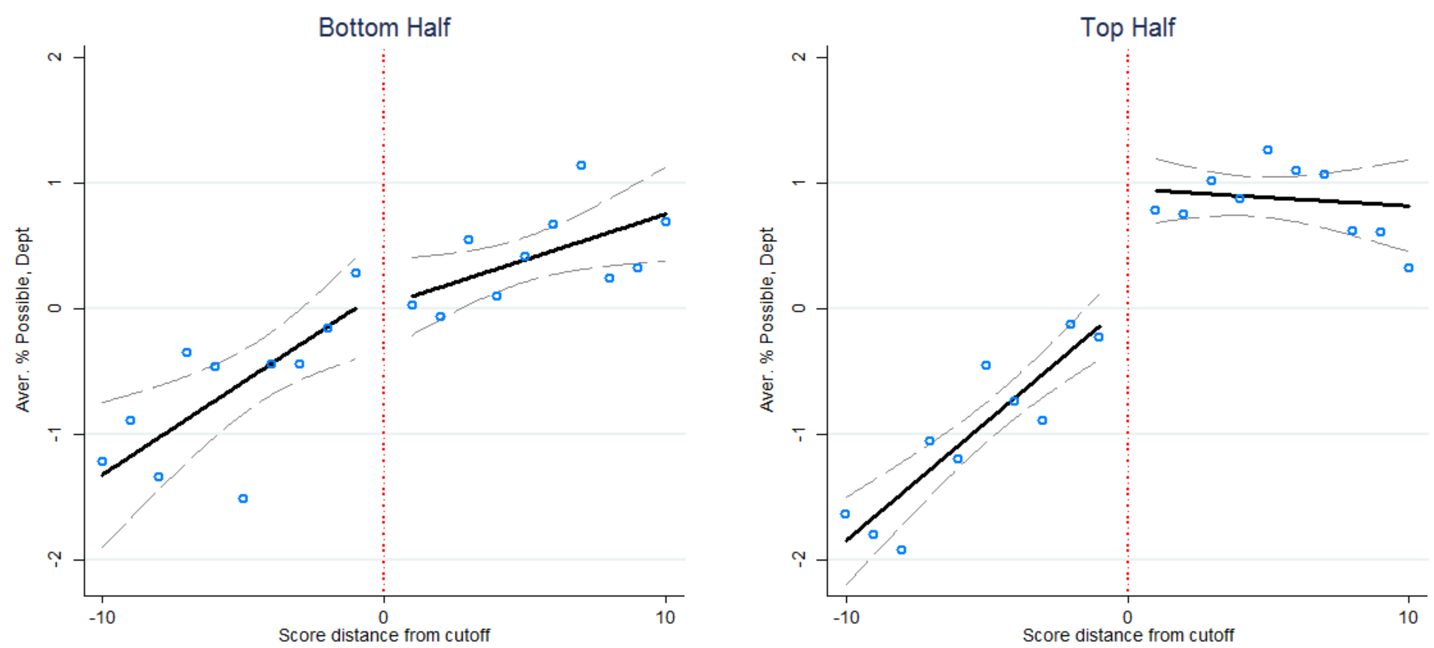
Notes: The three panels above give the smoothed fits of the log wages around the discontinuity. The bandwidth is 10 points on either side of the discontinuity. These figures display residuals are taken from a regression of the outcome on a linear trend in the test scores, including covariates and cut-off fixed effects. Only observations for that subgroup are used in the regression. Panel A (leftmost) uses the unconditional outcomes. Panel B (center) uses residuals from a regression on a linear trend in the test scores, including only cutoff fixed effects (i.e., no covariates). Panel C (rightmost) uses residuals from a regression on a linear trend in the test scores, including covariates and cutoff fixed effects. The dots in each panel represent the within-bin averages for each test point for that specific outcome and specification. Panel C corresponds what is shown in Table 5 column (11).

Figure 1.11: Non-linearity in Peer Quality: First-stage: Peer Quality at the College Level



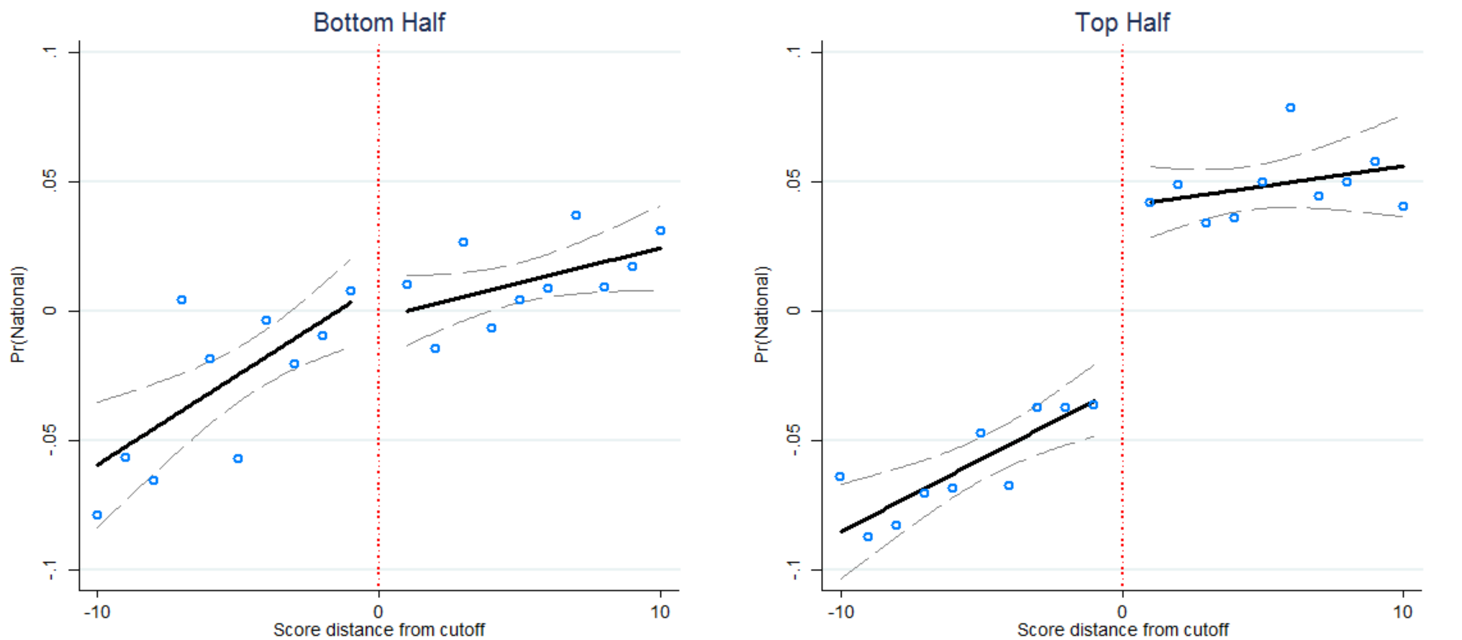
Notes: The two panels above give the smoothed fits of peer quality at the academic department level around the cutoff. The sample is split into 2 subgroups by whether the student applied to a department that was in the top-half or bottom-half of the peer quality distribution within that college major. Panel A (left) provides the results for those who applied to the bottom half in peer quality. Panel B (right) provides the results for those who applied to the top half. The bandwidth is 10 points on either side of the discontinuity. These figures display residuals are taken from a regression of the outcome on a linear trend in the test scores, including covariates and cut-off fixed effects. Only observations for that subgroup are used in the regression. These results correspond to coefficients in Table 8, Columns (1) and (2).

Figure 1.12: Non-linearity in Peer Quality: First-stage: Peer Quality at the Academic Department Level



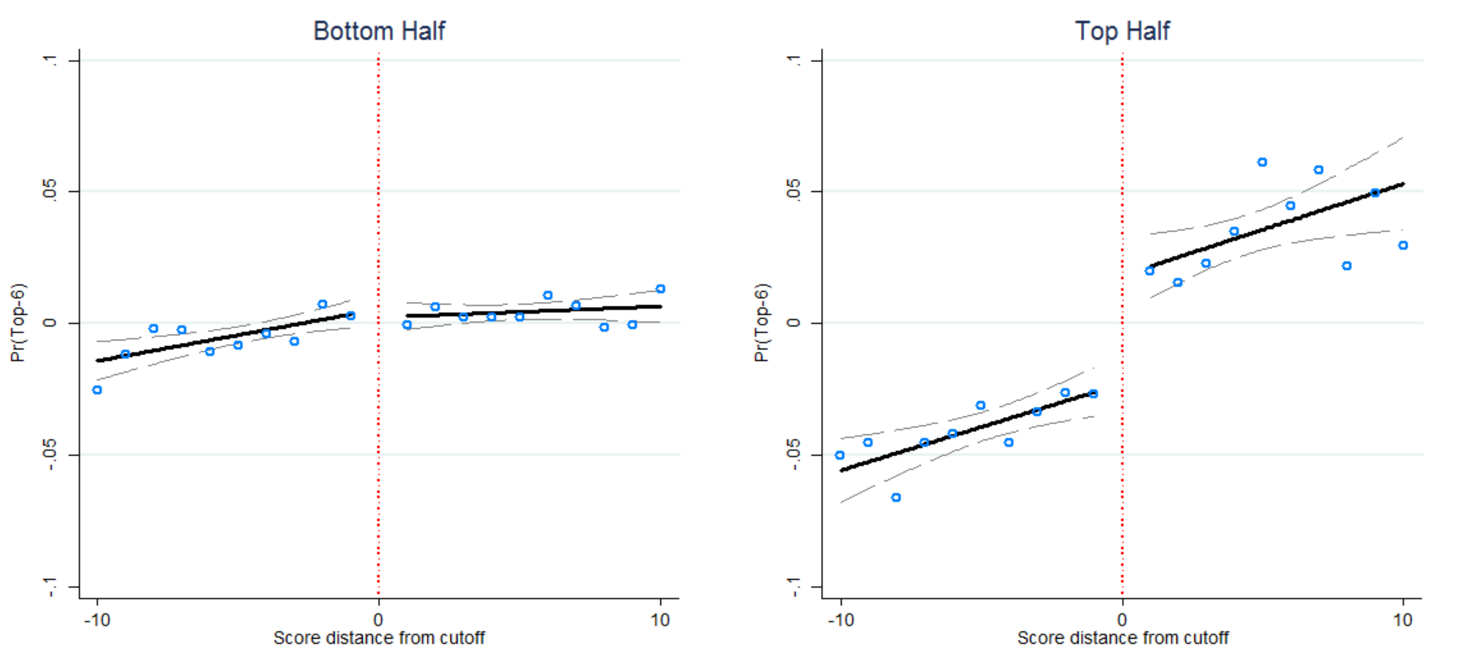
Notes: The two panels above give the smoothed fits of peer quality at the college level around the cutoff. The sample is split into 2 subgroups by whether the student applied to a department that was in the top-half or bottom-half of the peer quality distribution within that college major. Panel A (left) provides the results for those who applied to the bottom half in peer quality. Panel B (right) provides the results for those who applied to the top half. The bandwidth is 10 points on either side of the discontinuity. These figures display residuals are taken from a regression of the outcome on a linear trend in the test scores, including covariates and cut-off fixed effects. Only observations for that subgroup are used in the regression. These results correspond to coefficients in Table 8, Columns (3) and (4).

Figure 1.13: Non-linearity in Peer Quality: First-stage: Probability of Being Admitted to a National College.



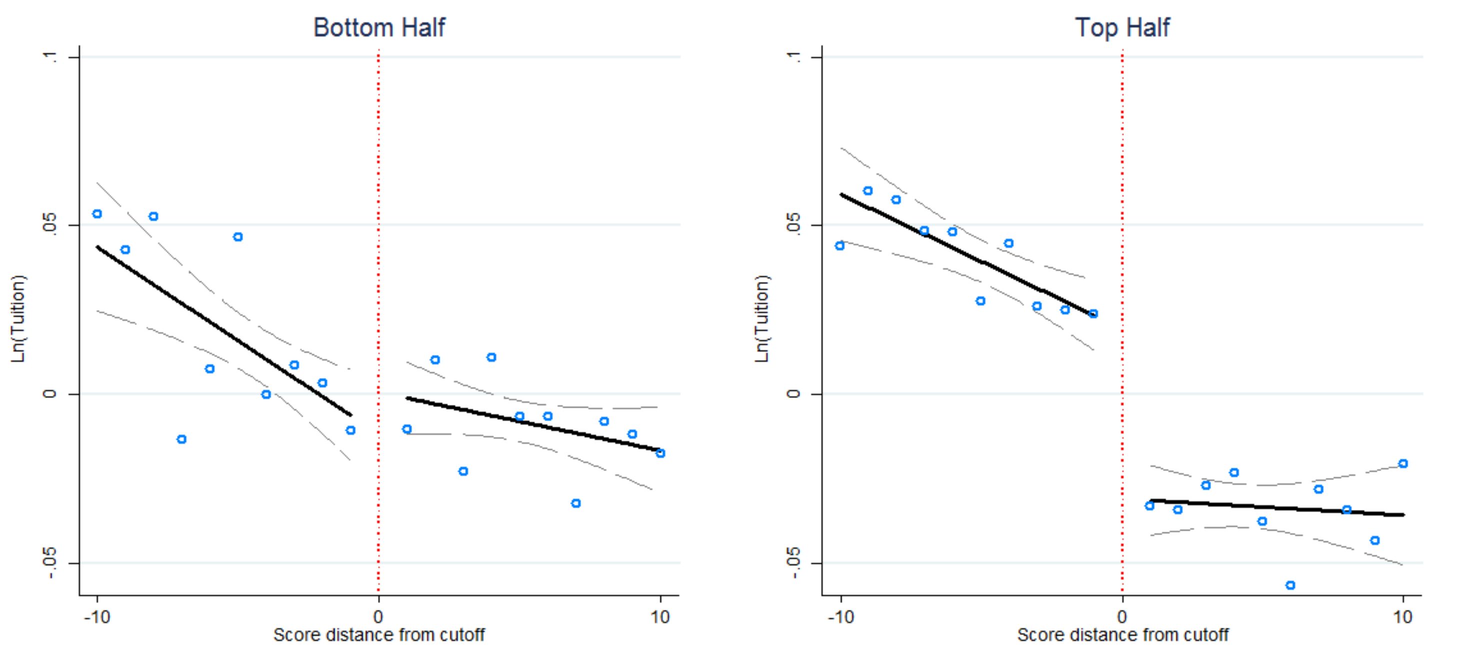
Notes: The two panels above give the smoothed fits of the probability of being admitted into a national college around the cutoff. The sample is split into 2 subgroups by whether the student applied to a department that was in the top-half or bottom-half of the peer quality distribution within that college major. Panel A (left) provides the results for those who applied to the bottom half in peer quality. Panel B (right) provides the results for those who applied to the top half. The bandwidth is 10 points on either side of the discontinuity. These figures display residuals are taken from a regression of the outcome on a linear trend in the test scores, including covariates and cut-off fixed effects. Only observations for that subgroup are used in the regression. These results correspond to coefficients in Table 8, Columns (5) and (6).

Figure 1.14: Non-linearity in Peer Quality: First-stage: Probability of Being Admitted to a Top-6 (Elite) National College.



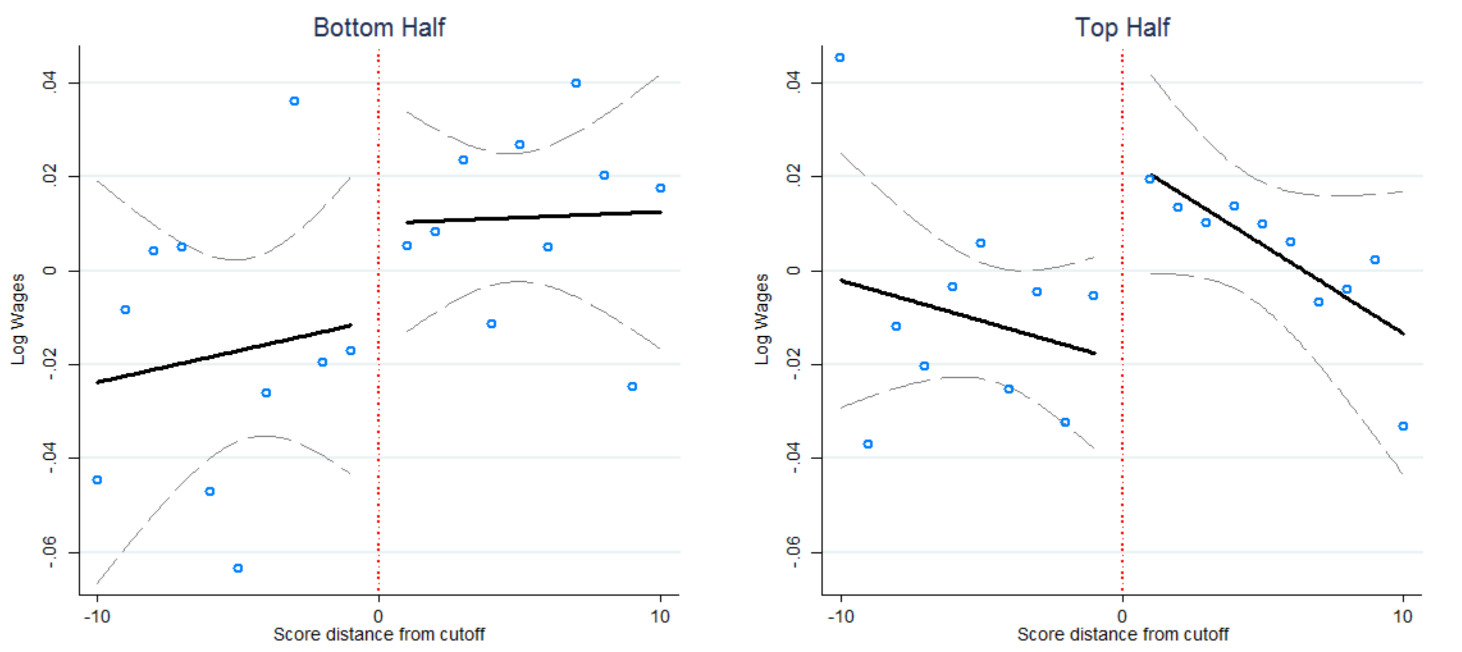
Notes: The two panels above give the smoothed fits of the probability of being admitted into a top-6 elite national college around the cutoff. The sample is split into 2 subgroups by whether the student applied to a department that was in the top-half or bottom-half of the peer quality distribution within that college major. Panel A (left) provides the results for those who applied to the bottom half in peer quality. Panel B (right) provides the results for those who applied to the top half. The bandwidth is 10 points on either side of the discontinuity. These figures display residuals are taken from a regression of the outcome on a linear trend in the test scores, including covariates and cut-off fixed effects. Only observations for that subgroup are used in the regression. These results correspond to coefficients in Table 8, Columns (7) and (8).

Figure 1.15: Non-linearity in Peer Quality: Reduced-Form: Log of Yearly Tuition.



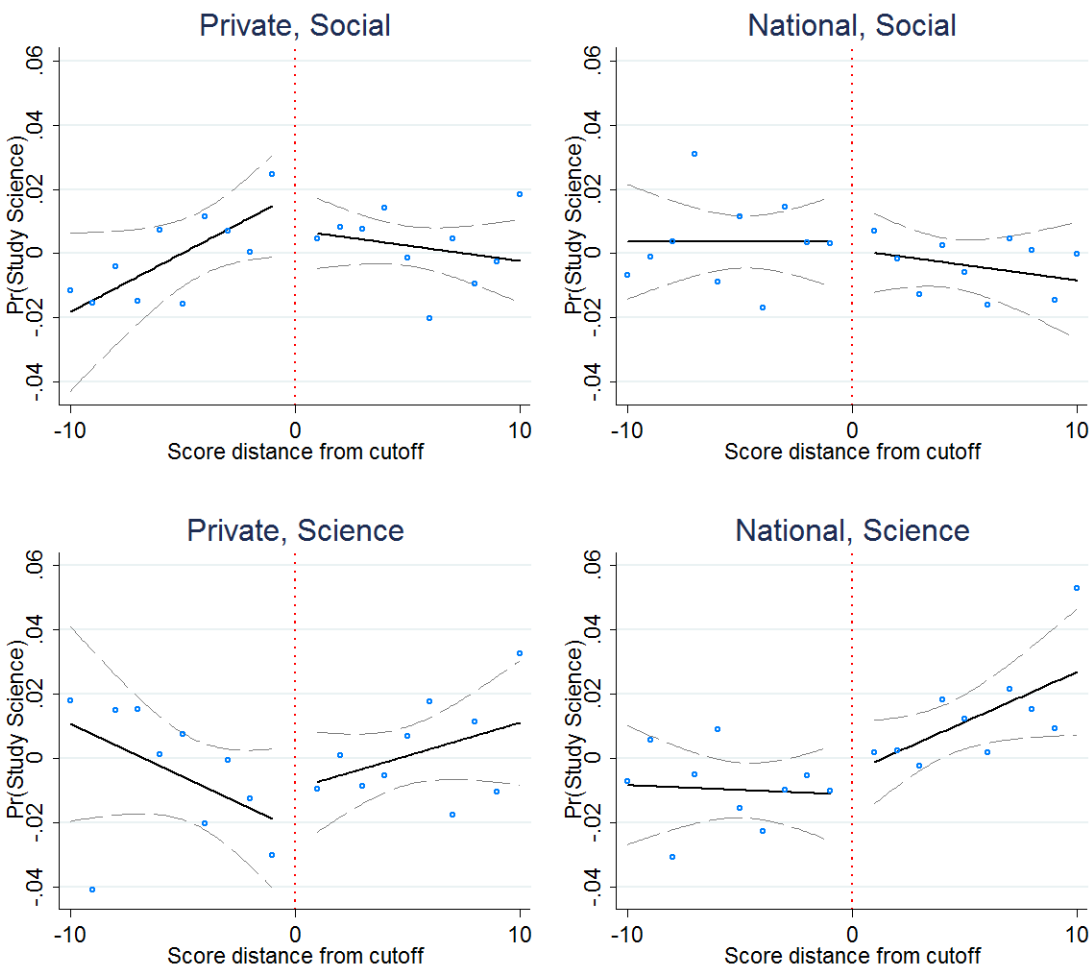
Notes: The two panels above give the smoothed fits of the log of yearly tuition around the cutoff. The sample is split into 2 subgroups by whether the student applied to a department that was in the top-half or bottom-half of the peer quality distribution within that college major. Panel A (left) provides the results for those who applied to the bottom half in peer quality. Panel B (right) provides the results for those who applied to the top half. The bandwidth is 10 points on either side of the discontinuity. These figures display residuals are taken from a regression of the outcome on a linear trend in the test scores, including covariates and cut-off fixed effects. Only observations for that subgroup are used in the regression. These results correspond to coefficients in Table 8, Columns (9) and (10).

Figure 1.16: Non-linearity in Peer Quality: Reduced-Form: Log Wages.



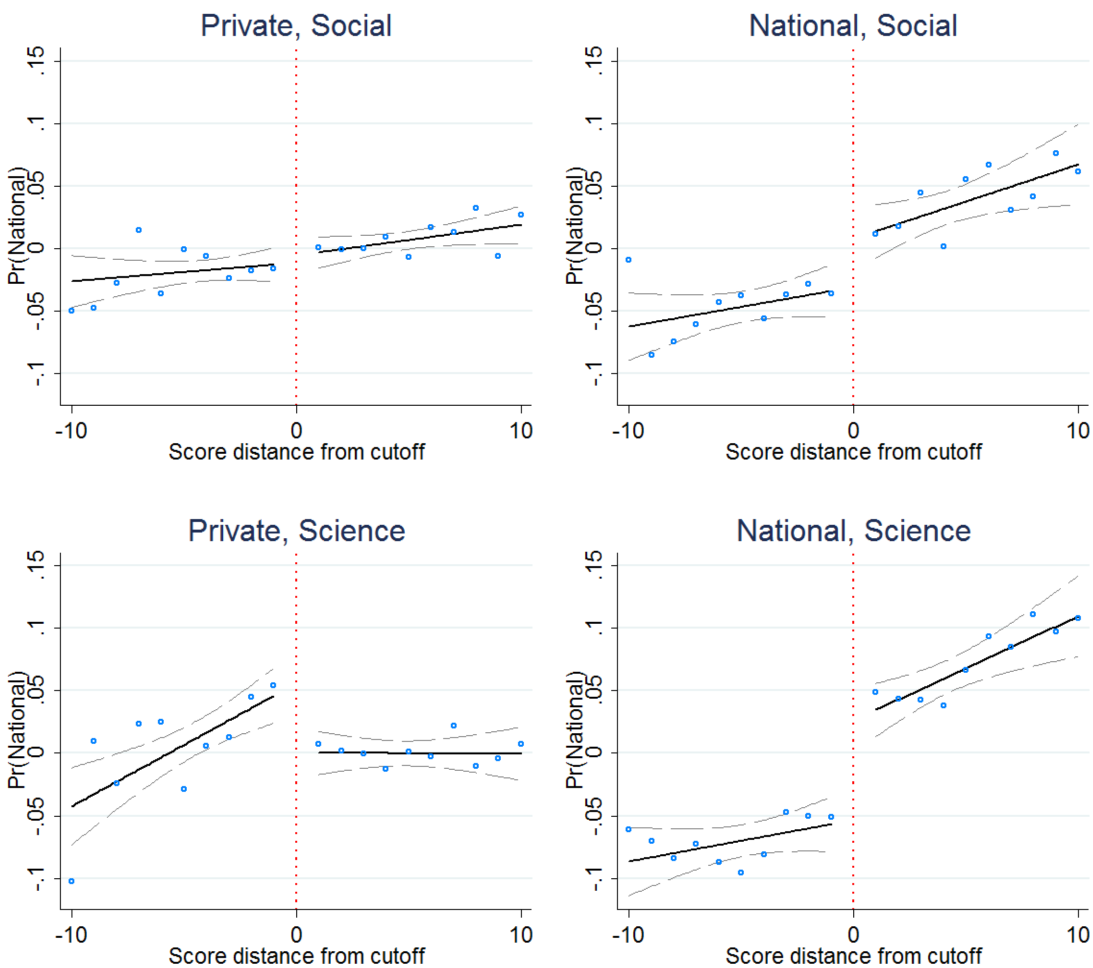
Notes: The two panels above give the smoothed fits of log-wages around the cutoff. The sample is split into 2 subgroups by whether the student applied to a department that was in the top-half or bottom-half of the peer quality distribution within that college major. Panel A (left) provides the results for those who applied to the bottom half in peer quality. Panel B (right) provides the results for those who applied to the top half. The bandwidth is 10 points on either side of the discontinuity. These figures display residuals are taken from a regression of the outcome on a linear trend in the test scores, including covariates and cut-off fixed effects. Only observations for that subgroup are used in the regression. These results correspond to coefficients in Table 8, Columns (11) and (12).

Figure 1.17: Heterogeneity in split samples: Probability of being admitted into the science track.



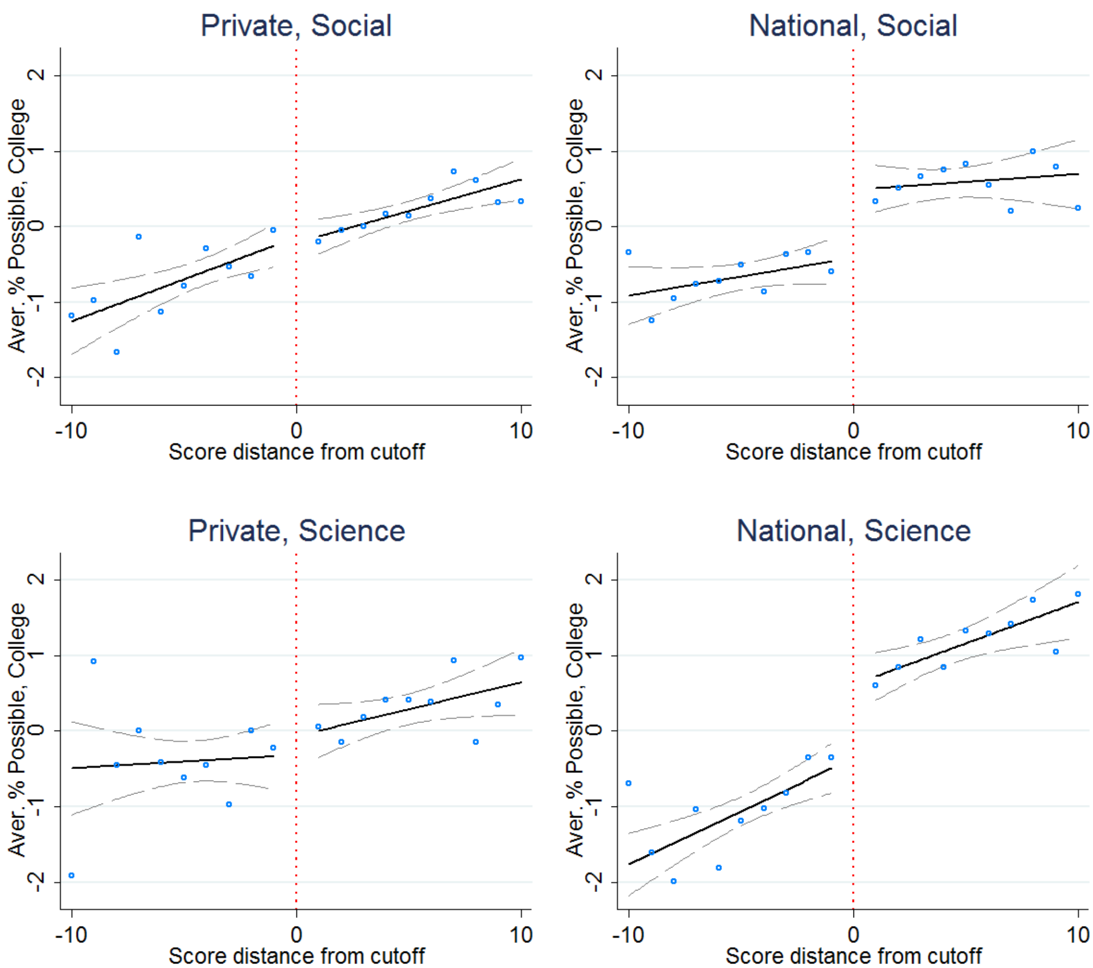
Notes: The four panels above give the smoothed fits of the probability of being admitted to a science-related college major around the cutoff. The sample is split into 4 groups by whether the student applied to a national versus a private college and to the social-track versus the science-track. Panel A (upper-left) provides the results for those who applied to the social track of a private college. Panel B (upper-right) provides the results for those who applied to the social track of a national college. Panel C (lower-left) provides the results for those who applied to the science track of a private college. Panel D (lower-right) provides the results for those who applied to the science track of a national college. The bandwidth is 10 points on either side of the discontinuity. These figures display residuals are taken from a regression of the outcome on a linear trend in the test scores, including covariates and cut-off fixed effects. Only observations for that subgroup are used in the regression.

Figure 1.18: Heterogeneity in split samples: National versus private college admissions.



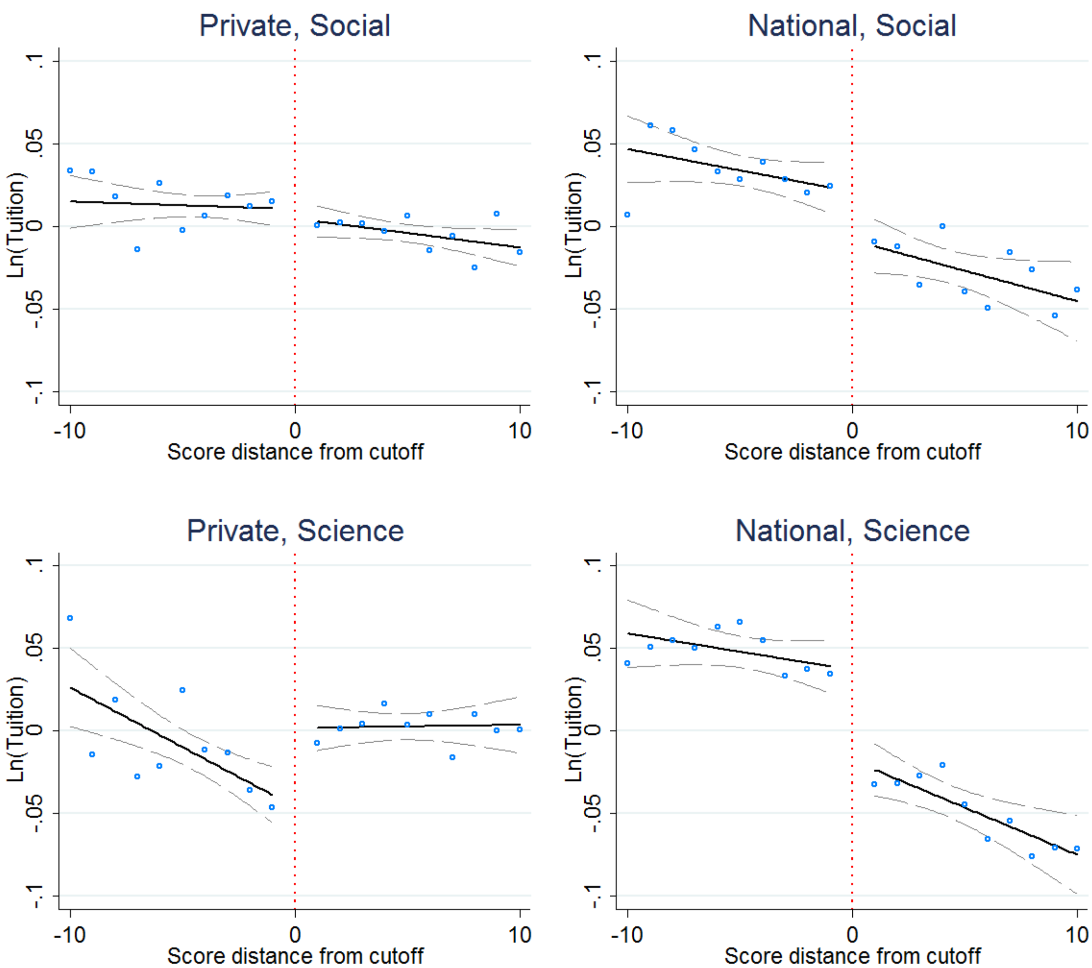
Notes: The four panels above give the smoothed fits of the probability of being admitted to a national college around the cutoff. The sample is split into 4 groups by whether the student applied to a national versus a private college and to the social-track versus the science-track. Panel A (upper-left) provides the results for those who applied to the social track of a private college. Panel B (upper-right) provides the results for those who applied to the social track of a national college. Panel C (lower-left) provides the results for those who applied to the science track of a private college. Panel D (lower-right) provides the results for those who applied to the science track of a national college. The bandwidth is 10 points on either side of the discontinuity. These figures display residuals are taken from a regression of the outcome on a linear trend in the test scores, including covariates and cut-off fixed effects. Only observations for that subgroup are used in the regression. Outcomes correspond to Table 9, Panel C.

Figure 1.19: Heterogeneity in split samples: Peer quality of the college.



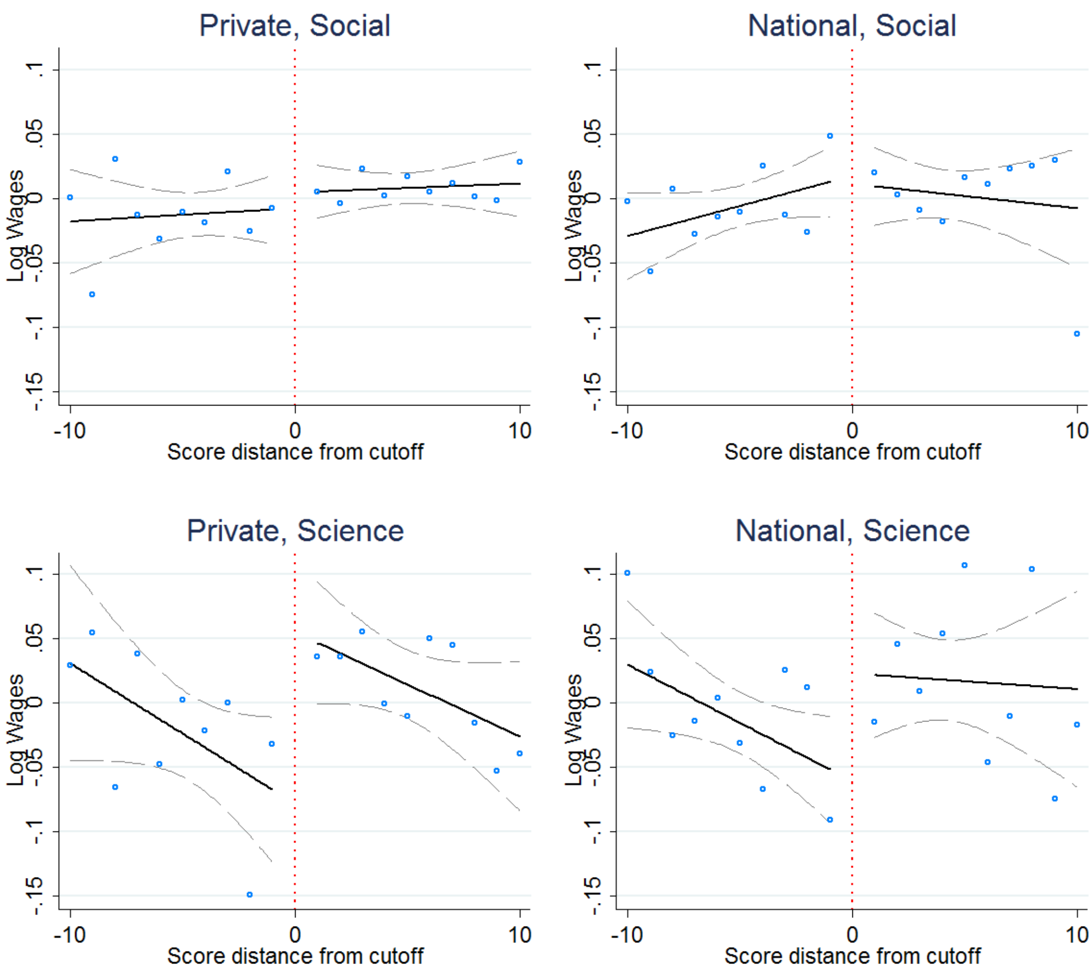
Notes: The four panels above give the smoothed fits of college peer quality around the cutoff. The sample is split into 4 groups by whether the student applied to a national versus a private college and to the social-track versus the science-track. Panel A (upper-left) provides the results for those who applied to the social track of a private college. Panel B (upper-right) provides the results for those who applied to the social track of a national college. Panel C (lower-left) provides the results for those who applied to the science track of a private college. Panel D (lower-right) provides the results for those who applied to the science track of a national college. The bandwidth is 10 points on either side of the discontinuity. These figures display residuals are taken from a regression of the outcome on a linear trend in the test scores, including covariates and cut-off fixed effects. Only observations for that subgroup are used in the regression. Outcomes correspond to Table 9, Panel A.

Figure 1.20: Heterogeneity in split samples: Log yearly tuition.



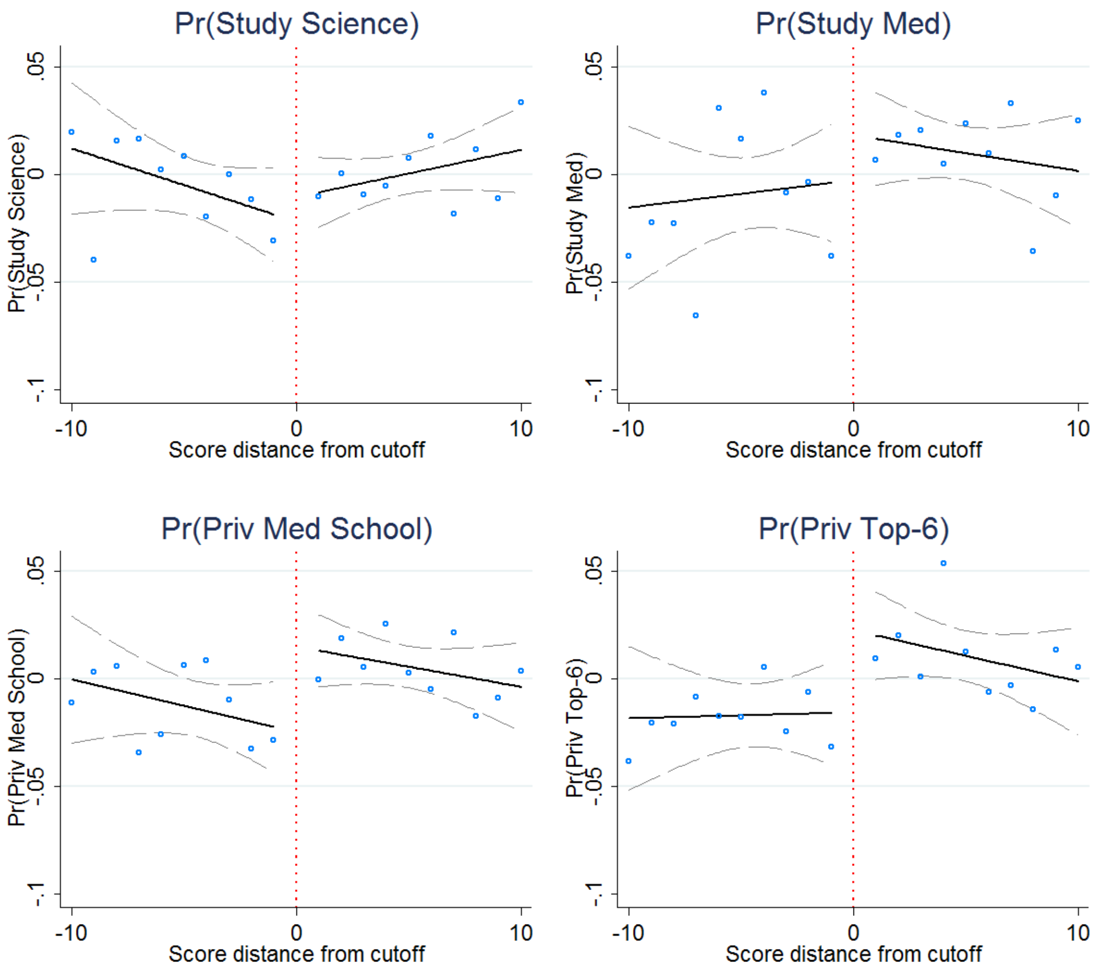
Notes: The 4 panels above give the smoothed fits of the log of yearly tuition around the cutoff. The sample is split into 4 groups by whether the student applied to a national versus a private college and to the social-track versus the science-track. Panel A (upper-left) provides the results for those who applied to the social track of a private college. Panel B (upper-right) provides the results for those who applied to the social track of a national college. Panel C (lower-left) provides the results for those who applied to the science track of a private college. Panel D (lower-right) provides the results for those who applied to the science track of a national college. The bandwidth is 10 points on either side of the discontinuity. These figures display residuals are taken from a regression of the outcome on a linear trend in the test scores, including covariates and cut-off fixed effects. Only observations for that subgroup are used in the regression. Outcomes correspond to Table 9, Panel E.

Figure 1.21: Heterogeneity in split samples: Log wages.



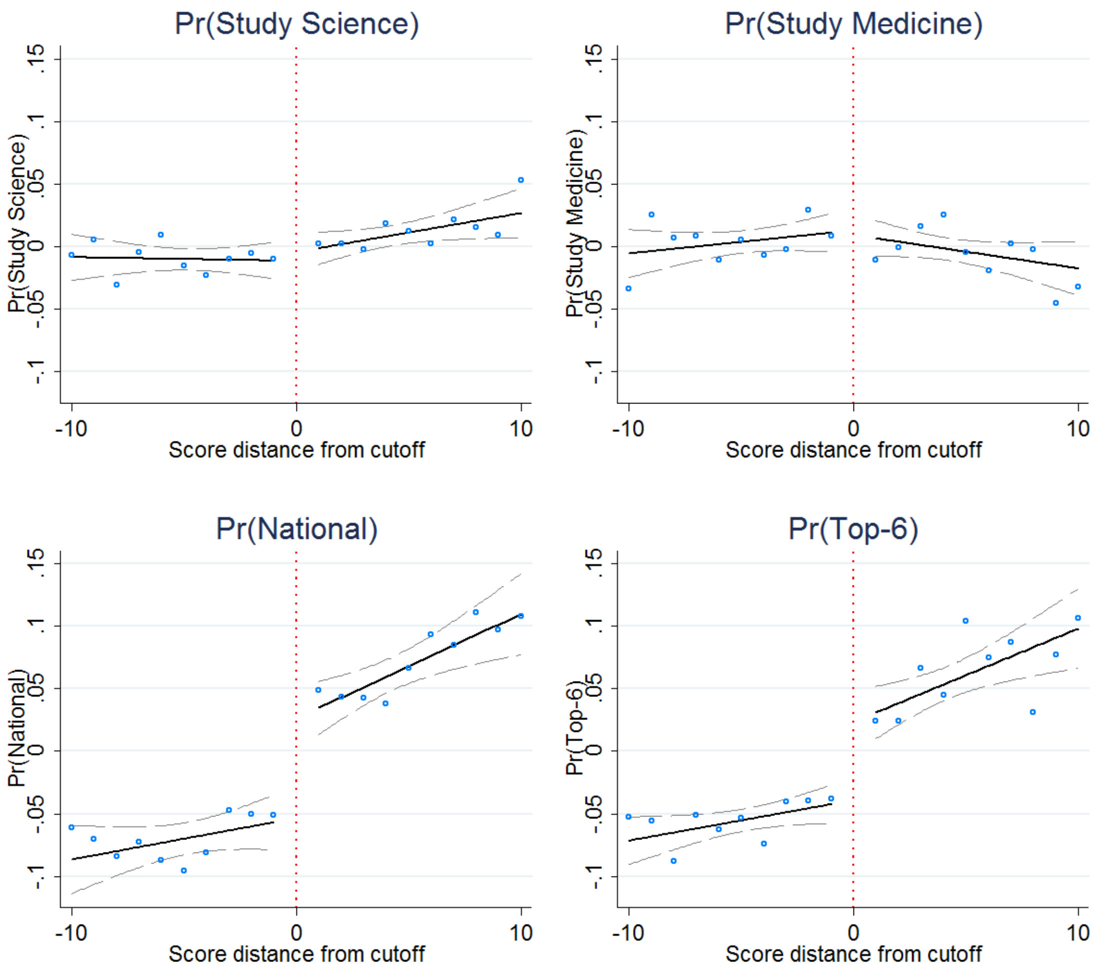
Notes: The 4 panels above give the smoothed fits of log wages around the cutoff. The sample is split into 4 groups by whether the student applied to a national versus a private college and to the social-track versus the science-track. Panel A (upper-left) provides the results for those who applied to the social track of a private college. Panel B (upper-right) provides the results for those who applied to the social track of a national college. Panel C (lower-left) provides the results for those who applied to the science track of a private college. Panel D (lower-right) provides the results for those who applied to the science track of a national college. The bandwidth is 10 points on either side of the discontinuity. These figures display residuals are taken from a regression of the outcome on a linear trend in the test scores, including covariates and cut-off fixed effects. Only observations for that subgroup are used in the regression. Outcomes correspond to Table 9, Panel F.

Figure 1.22: Heterogeneity for Private-Science Programs.



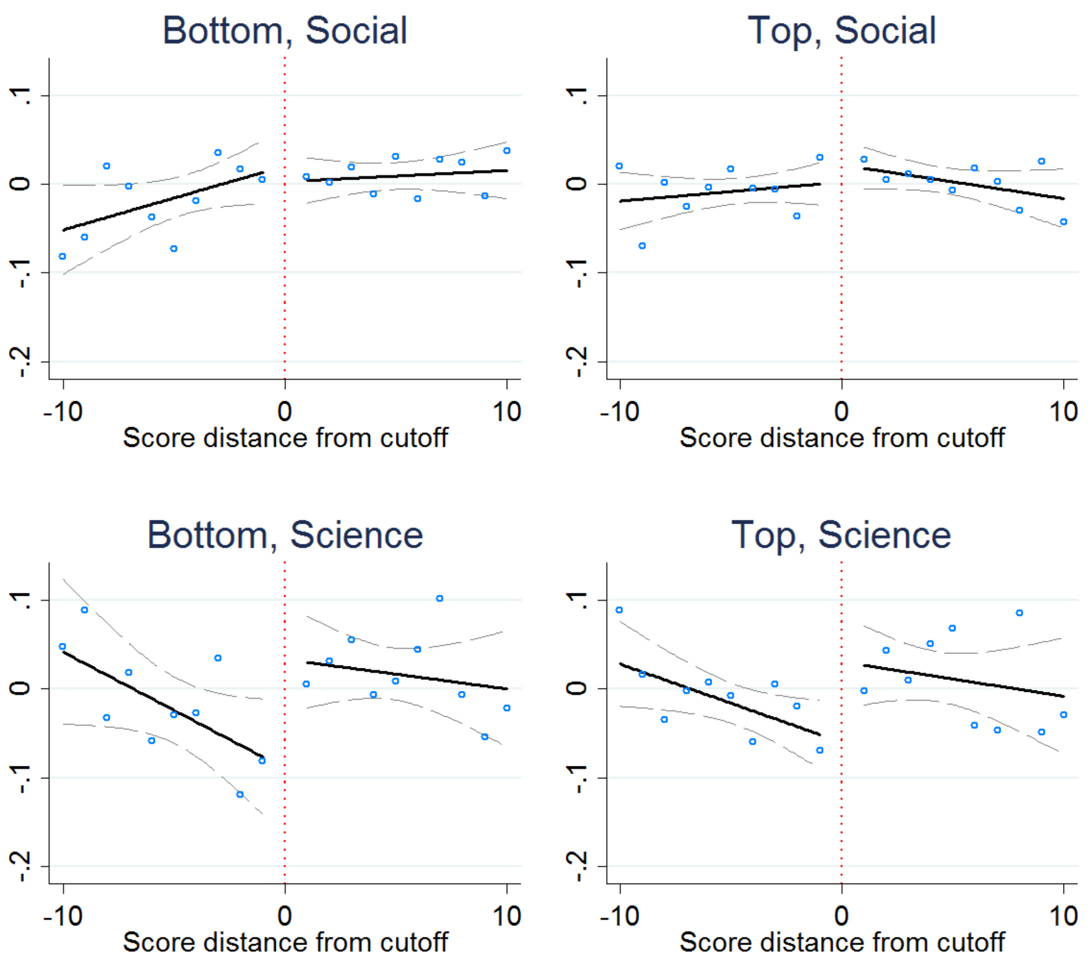
Notes: The 4 panels above give the smoothed fits of for four separate outcomes only for the subgroup of private-science applicants. Panel A (upper-left) provides the results for the probability of being admitted at all into a science-track (vs. social-track) college major. Panel B (upper-right) provides the results for the probability of being admitted at all into a medical-science college major only versus not. Panel C (lower-left) provides the results for the probability of being admitted at all into a medical school versus not. Panel D (lower-right) provides the results for the probability of being admitted at all into a top-6 private college versus not. The bandwidth is 10 points on either side of the discontinuity. Outcomes correspond to Table 10, Panel A.

Figure 1.23: Heterogeneity for National-Science Programs.



Notes: The 4 panels above give the smoothed fits of for four separate outcomes only for the subgroup of national-science applicants. Panel A (upper-left) provides the results for the probability of being admitted at all into a science-track (vs. social-track) college major. Panel B (upper-right) provides the results for the probability of being admitted at all into a medical-science college major only versus not. Panel C (lower-left) provides the results for the probability of being admitted at all into a national college versus not. Panel D (lower-right) provides the results for the probability of being admitted at all into a top-6 national college versus not. The bandwidth is 10 points on either side of the discontinuity. Outcomes correspond to Table 10, Panel B.

Figure 1.24: Heterogeneity for top- vs. bottom-half peer quality and college major: Log wages.



Notes: The 4 panels above give the smoothed fits of log wages around the cutoff. The sample is split into 4 groups by whether the student applied to an academic department that was either at the top-half versus the bottom-half of the peer quality distribution within their college major and to the social-track versus the science-track. Panel A (upper-left) provides the results for those who applied to the social track of a bottom-half department. Panel B (upper-right) provides the results for those who applied to the social track of a top-half department. Panel C (lower-left) provides the results for those who applied to the science track of a bottom-half department. Panel D (lower-right) provides the results for those who applied to the science track of a top-half department. The bandwidth is 10 points on either side of the discontinuity. These figures display residuals are taken from a regression of the outcome on a linear trend in the test scores, including covariates and cut-off fixed effects. Only observations for that subgroup are used in the regression.

Tables

Table 1.1: Descriptive Statistics

Variables	Obs	Mean	Std. Dev.	Min	Max
A. First-stage outcomes					
Pr(National)	36483	0.32	0.46	0	1
Pr(Top-6)	36483	0.10	0.30	0	1
Log Tuition	36483	10.53	0.34	9.89	11.11
B. Reduced-form outcomes					
Log Wages	11337	10.01	0.35	9.31	10.65
C. Student covariates					
Male	36483	0.53	0.50	0	1
Age	36483	18.38	0.42	16.67	21.67
Father's Education	36455	11.22	3.43	0	18
Mother's Education	36455	10.06	3.47	0	18
Log of Family Income	31492	10.63	0.48	9.31	11.79

Table 1.2: Check for Balance of Covariates.

	Have Wage (1)	Pr(Male) (2)	Age (3)	Mother's YOS (4)	Father's YOS (5)	ln(Parental Income) (6)
1(Treat)	0.014	-0.015+	-0.001	0.114+	0.004	0.004
	(0.013)	(0.008)	(0.009)	(0.059)	(0.046)	(0.013)
R2	0.22	0.03	0.02	0.43	0.43	0.14
N	24885	24885	24885	24885	24885	24885

Notes: The six student characteristics above are used to show that the treated (above cutoff) students are generally comparable to their counterfactuals in the control (below cutoff) group. Each coefficient in Columns (1) through (6) is the output from regressing one covariate on a treatment indicator, a linear trend in the running variable, and the other covariates and fixed effects. The bandwidth is set to 10 points. Column (1) shows balance in the probability that a student for which I have test data will show up in my earnings data. Column (2) shows balance in the proportion male. Column (3) shows balance in the age of students. Columns (4) and (5) show the balance in mother and father's years of schooling at time of student's birth. (Mothers of children slightly just above the cutoff appear to have slightly more schooling, significant at the 10% alpha level.) Column (6) shows the balance in parental income at time of application. Standard errors in parentheses. + $p<0.10$, * $p<0.05$, *** $p<0.01$

Table 1.3: Cross-validated first-stage results on peer quality at college and department level.

	A. Peer Quality at College						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
1(Treat)	1.737***	0.306*	0.549***	0.716***	0.924***	1.111***	1.242***
	(0.146)	(0.088)	(0.114)	(0.129)	(0.151)	(0.188)	(0.206)
bwidth	1	4	7	10	13	16	19
R2	0.50	0.50	0.50	0.49	0.49	0.49	0.49
N	3949	15141	23394	28571	31283	32522	33023
	B. Peer Quality at Department						
	(8)	(9)	(10)	(11)	(12)	(13)	(14)
1(Treat)	1.692***	0.059	0.256+	0.464*	0.779***	0.917***	1.077***
	(0.074)	(0.141)	(0.153)	(0.181)	(0.223)	(0.249)	(0.274)
bwidth	1	4	7	10	13	16	19
R2	0.43	0.43	0.42	0.42	0.42	0.42	0.43
N	3949	15141	23394	28571	31283	32522	33023

Notes: Regressions use a linear specification of the model. Cutoff fixed effects are included. Covariates are included. Bandwidth ranges from 1 to 19 points around the cutoff, with the IK bandwidth just over 4. Robust standard errors are clustered on the test-point level (the running variable). Standard errors in parentheses. + p<0.10, * p<0.05, *** p<0.01

Table 1.4: Cross-validated first-stage results on prestige effects.

	A. Probability of being admitted into a national (public) college						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
1(Treat)	0.102***	0.038***	0.038***	0.043***	0.056***	0.066***	0.074***
	(0.007)	(0.007)	(0.005)	(0.005)	(0.007)	(0.009)	(0.010)
	14.208	5.442	7.296	8.740	7.833	7.673	7.420
bwidth	1	4	7	10	13	16	19
R2	0.36	0.35	0.35	0.34	0.34	0.34	0.34
N	3949	15141	23394	28571	31283	32522	33023
	B. Probability of being admitted into a top-6 (elite public) college						
	(8)	(9)	(10)	(11)	(12)	(13)	(14)
1(Treat)	0.057*	0.017*	0.023***	0.033***	0.046***	0.056***	0.061***
	(0.006)	(0.006)	(0.004)	(0.005)	(0.008)	(0.010)	(0.011)
	9.612	3.109	6.151	6.233	5.593	5.658	5.763
bwidth	1	4	7	10	13	16	19
R2	0.31	0.29	0.29	0.28	0.27	0.27	0.27
N	3949	15141	23394	28571	31283	32522	33023

Notes: Regressions use a linear specification of the model. Cutoff fixed effects are included. Covariates are included. Bandwidth ranges from 1 to 19 points around the cutoff, with the IK bandwidth just over 4. Robust standard errors are clustered on the test-point level (the running variable). Standard errors in parentheses. + p<0.10, * p<0.05, *** p<0.01

Table 1.5: Cross-validated monetary effects.

	A. Log Yearly Tuition						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
1(Treat)	-0.072***	-0.030***	-0.026***	-0.031***	-0.040***	-0.046***	-0.052***
	(0.005)	(0.005)	(0.005)	(0.004)	(0.006)	(0.006)	(0.007)
bwidth?	1	4	7	10	13	16	19
R2	0.29	0.29	0.28	0.28	0.28	0.28	0.28
N	3949	15141	23394	28571	31283	32522	33023
	B. Log Wages						
	(8)	(9)	(10)	(11)	(12)	(13)	(14)
1(Treat)	0.027***	0.026*	0.020*	0.032***	0.028*	0.031***	0.035***
	(0.005)	(0.011)	(0.009)	(0.011)	(0.010)	(0.009)	(0.009)
bwidth?	1	4	7	10	13	16	19
R2	0.18	0.14	0.13	0.12	0.12	0.12	0.12
N	1274	4670	7226	8864	9712	10116	10287

Notes: Regressions use a linear specification of the model. Cutoff fixed effects are included. Covariates are included. Bandwidth ranges from 1 to 19 points around the cutoff, with the IK bandwidth just over 4. Robust standard errors are clustered on the test-point level (the running variable). Standard errors in parentheses. + $p<0.10$, * $p<0.05$, *** $p<0.01$

Table 1.6: Cross-validated wage effects, with and without covariates.

	A. With covariates						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
1(Treat)	0.027***	0.026*	0.020*	0.032***	0.028*	0.031***	0.035***
	(0.005)	(0.011)	(0.009)	(0.011)	(0.010)	(0.009)	(0.009)
bwidth?	1	4	7	10	13	16	19
R2	0.18	0.14	0.13	0.12	0.12	0.12	0.12
N	1274	4670	7226	8864	9712	10116	10287
	B. Without covariates						
	(8)	(9)	(10)	(11)	(12)	(13)	(14)
1(Treat)	0.033***	0.037***	0.030***	0.040***	0.037***	0.040***	0.042***
	(0.006)	(0.014)	(0.012)	(0.013)	(0.012)	(0.011)	(0.010)
bwidth?	1	4	7	10	13	16	19
R2	0.09	0.05	0.03	0.03	0.03	0.03	0.03
N	1275	4674	7235	8876	9726	10131	10303

Notes: The regressions replicate the monetary effects in Table 5, Panel B. Regressions use a linear specification of the model. Cutoff fixed effects are included. Covariates are included in Panel A and not included in Panel B. Bandwidth ranges from 1 to 19 points around the cutoff, with the IK bandwidth just over 4. Robust standard errors are clustered on the test-point level (the running variable). Standard errors in parentheses. + $p<0.10$, * $p<0.05$, *** $p<0.01$

Table 1.7: Cross-validated wage effects, by gender.

	A. Females						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
1(Treat)	0.028***	0.031*	0.010	0.019+	0.018+	0.024*	0.028***
	(0.009)	(0.011)	(0.011)	(0.012)	(0.011)	(0.011)	(0.010)
bwidth?	1	4	7	10	13	16	19
R2	0.12	0.06	0.05	0.05	0.04	0.05	0.04
N	1012	3689	5697	6977	7656	7965	8095
	B. Males						
	(8)	(9)	(10)	(11)	(12)	(13)	(14)
1(Treat)	0.016	0.016	0.057*	0.069*	0.053*	0.049*	0.052***
	(0.012)	(0.028)	(0.029)	(0.026)	(0.024)	(0.021)	(0.020)
bwidth?	1	4	7	10	13	16	19
R2	0.17	0.06	0.05	0.04	0.04	0.04	0.04
N	262	981	1529	1887	2056	2151	2192

Notes: The regressions replicate the monetary effects in Table 5, Panel B for male versus female by running the same regressions splitting into gender-specific subsamples. Regressions use a linear specification of the model. Cutoff fixed effects are included. Covariates are included. Bandwidth ranges from 1 to 19 points around the cutoff, with the IK bandwidth just over 4. Robust standard errors are clustered on the test-point level (the running variable). Standard errors in parentheses. + $p<0.10$, * $p<0.05$, *** $p<0.01$

Table 1.8: Nonlinearity in peer quality: Outcomes by top- versus bottom-half department.

	A. Peer @ College		B. Peer @ Dept	
	(1)	(2)	(3)	(4)
1(Treat)	0.365	1.190***	0.204	1.073***
	-0.287	-0.176	-0.279	-0.213
Half?	Bottom	Top	Bottom	Top
R2	0.31	0.46	0.29	0.41
N	10172	19700	10172	19700
	C. Pr(National)		D. Pr(Top-6)	
	(5)	(6)	(7)	(8)
1(Treat)	0.002	0.078***	-0.003	0.051***
	-0.012	-0.008	-0.003	-0.008
Half?	Bottom	Top	Bottom	Top
R2	0.2	0.33	0.13	0.27
N	10172	19700	10172	19700
	E. Ln(Tuition)		F. Ln(Wages)	
	(9)	(10)	(11)	(12)
1(Treat)	0.005	-0.055***	0.018	0.047***
	-0.008	-0.005	-0.015	-0.012
Half?	Bottom	Top	Bottom	Top
R2	0.17	0.27	0.11	0.13
N	10172	19700	3562	5728

Notes: For each outcome above, there are two columns corresponding to the coefficients for two subgroups. These subgroups are formed based on where the student applied early, whether it was an academic department that was in the top-half or bottom-half of the peer quality distribution within that college major. For each outcome, "Bottom" signifies that the coefficient is for those who applied to the bottom half in peer quality. For each outcome, "Top" provides the results for those who applied to the top half. The bandwidth is 10 points on either side of the discontinuity. The regression follows previous ones, including covariates and cutoff fixed-effects and clustering robust standard errors at the test-point level in the running variables. These results correspond to Figures 11 through 16. Standard errors in parentheses. + $p<0.10$, * $p<0.05$, *** $p<0.01$

Table 1.9: Nonlinearities in college rank and major.

	A. Peer @ College				B. Peer @ Dept			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
1(Treat)	0.286	0.864*	1.122***	1.204***	0.510	0.910*	0.277	1.370***
	(0.280)	(0.377)	(0.231)	(0.264)	(0.326)	(0.394)	(0.182)	(0.376)
Rank?	Private	Private	Nat'l	Nat'l	Private	Private	Nat'l	Nat'l
Track?	Social	Science	Social	Science	Social	Science	Social	Science
R2	0.24	0.44	0.43	0.41	0.36	0.49	0.38	0.50
N	8232	4955	9017	8980	8232	4955	9017	8980
	C. Pr(National)				D. Pr(Top-6)			
	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
1(Treat)	0.011	-0.040***	0.065***	0.105***	0.004*	-0.011	0.039*	0.068***
	(0.007)	(0.013)	(0.020)	(0.021)	(0.001)	(0.007)	(0.014)	(0.011)
Rank?	Private	Private	Nat'l	Nat'l	Private	Private	Nat'l	Nat'l
Track?	Social	Science	Social	Science	Social	Science	Social	Science
R2	0.09	0.25	0.29	0.26	0.03	0.28	0.22	0.28
N	8232	4955	9017	8980	8232	4955	9017	8980
	E. Ln(Tuition)				F. Ln(Wages)			
	(17)	(18)	(19)	(20)	(21)	(22)	(23)	(24)
1(Treat)	-0.010+	0.033***	-0.046***	-0.074***	0.018	0.125***	-0.007	0.084*
	(0.006)	(0.011)	(0.014)	(0.014)	(0.014)	(0.040)	(0.023)	(0.038)
Rank?	Private	Private	Nat'l	Nat'l	Private	Private	Nat'l	Nat'l
Track?	Social	Science	Social	Science	Social	Science	Social	Science
R2	0.07	0.20	0.28	0.24	0.09	0.21	0.10	0.18
N	8232	4955	9017	8980	3887	1149	3223	1513

Notes: For each outcome above, there are four columns corresponding to the coefficients for four subgroups. These subgroups are formed based on where the student applied, whether it was an academic department that was in a private versus national college or in social versus science tracks. For each outcome, "Private" and "Nat'l" signifies that the coefficient is for those who applied to private vs. national colleges. For each outcome, "Social" vs. "Science" provides the results for those who applied to a social- vs. science-track college major. The bandwidth is 10 points on either side of the discontinuity. The regression follows previous ones, including covariates and cutoff fixed-effects and clustering robust standard errors at the test-point level in the running variables. These results correspond to Figures 17 through 22. Standard errors in parentheses. + $p<0.10$, * $p<0.05$, *** $p<0.01$

Table 1.10: Nonlinearities in college major vs rank for science majors.

A. Private-Science

	Pr(Science) (1)	Pr(Med) (2)	Pr(Priv Med School) (3)	Pr(Private Top-6) (4)
1(Treat)	0.025+	0.027	0.042***	0.040***
	(0.013)	(0.018)	(0.008)	(0.012)
R2	0.12	0.23	0.15	0.12
N	4955	4955	4955	4955

B. National-Science

	Pr(Science) (5)	Pr(Med) (6)	Pr(National) (7)	Pr(Top-6) (8)
1(Treat)	0.013*	-0.011	0.105***	0.068***
	(0.006)	(0.010)	(0.021)	(0.011)
R2	0.07	0.12	0.26	0.28
N	8980	8980	8980	8980

Notes: For each outcome above, there are two columns corresponding to the coefficients for two subgroups. These subgroups are formed based on where the student applied, whether it was an academic department that was in the top-half or bottom-half of the peer quality distribution within that college major. For each outcome, "Bottom" signifies that the coefficient is for those who applied to the bottom half in peer quality. For each outcome, "Top" provides the results for those who applied to the top half. The bandwidth is 10 points on either side of the discontinuity. The regression follows previous ones, including covariates and cutoff fixed-effects and clustering robust standard errors at the test-point level in the running variables. These results correspond to Figures 23 and 24. Standard errors in parentheses. + $p<0.10$, * $p<0.05$, *** $p<0.01$

Chapter 2

Returns to College Selectivity: RD Evidence from a Centralized Admissions System

Attending a prestigious college is the goal of many aspiring students and their parents. It is assumed that prestige will serve as the catalyst to a positive and virtuous cycle of growth and prosperity. In this paper, therefore, I estimate the impact of going to a prestigious college on a later-life outcome, earnings.

Estimating the impact of college prestige on earnings is not a new endeavor (Chevalier and Conlon, 2003). Black and Smith (2006) helpfully reviews some of this literature, detailing the difficulty of precisely and convincingly estimating its influence, given the then current limitations of data and methodology. My paper addresses these limitations using data from Taiwan and a regression discontinuity (RD) design. Through these, I provide credible quasi-experimental evidence that, on average, college applicants for the same college major who score just above the admissions threshold (versus just below) to a national college are: a.) 50% more likely to get into a national college, b.) experience 0.2 σ increase in average peer quality, and c.) enjoy a 0.6 *decrease* in tuition costs. Long-term, this prestige advantage corresponds to an improvement of 0.07 log-points (on a 2-digit log base) or 0.22 σ in log earnings, or around 5% more every month. These results are economically substantive, statistically significant, and robust to specification testing.

My paper contributes to the 'college quality' literature in terms of data quantity and quality as well as research design. First, the data from Taiwan is comprehensive: I use a national universe of microdata for college examinations, college admissions, and earnings — all merged on national IDs — to identify causal estimates of the returns to college quality. Further, because my data catalogs student admissions down to the academic department level, I can estimate these returns at the college-major level, instead of just the college level, as in the other papers. Thus, my estimates can take into account the choice of college major as well as the choice of college, which had before been the purview of structural estimation strategies (Arcidiacono, 2004).

Second, the quality of the data is exceptional. Having the universe of data for this admissions cohort means I also have the universe of discontinuities. I estimate these reduced-forms results using an RD framework with discontinuities that are generated by the Taiwanese government's centralized student allocation system. Since I am estimating the returns to being admitted to a prestigious versus non-prestigious college, I will be exploiting only the discontinuities associated with the lowest score needed to be admitted into a national college within that specific college major. Thus, I use only a subset of the total number of discontinuities available to me.

Third, my research design employs a stacked RD; it allows me to approximate a local average treatment effect (LATE) at many different points in the test-score distribution, which gives greater representativeness to these results. Further, this research design allows me to answer the question of whether (in this context) it is better to be a big fish in a little pond or a little fish in a big pond. The results suggest the latter.

Following the recent college quality literature (Black and Smith, 2006; Dale and Krueger, 2002, 2011), I measure college quality using three common proxies. The most common proxy views college quality as a function of peer quality, proxied by the average college entrance exam score at that college. I utilize this, but interpret the results mainly through the second quality indicator, college prestige. Prestige in my study is proxied by the university's status as either a national or a private college (Chevalier and Conlon, 2003), which in the Taiwanese context is an important label of overall quality. Third, I show that financial aid in the form of public subsidy (Van der Klaauw, 2002; Winston, 1999), is highly correlated with being in a national college or university in Taiwan.[1]

Other mechanisms in addition to prestige may also be driving the reduced-form results I identify, such as peer group structure or tracking while in college (Carrell et al., 2011); yet

[1] Other measures of college quality have been considered in the literature: such as endowments (Winston, 1999), faculty quality and faculty-student ratios (Saavedra, 2008; Winston, 1999), or research productivity (Aghion et al., 2010). Each of these papers shows that peer quality is correlated with the other proxies of college quality, and thus my proxy is still a valid summary.

the positive wage effects I identify fall in line with what Wiske-Dillon and Smith (2012) — a structural estimation — find. Among college students there appears to be a preference for being a little fish in a big pond, or as Wiske-Dillon and Smith (2012) call it "being underqualified." In other words, students prefer being in a higher quality school, even if they are among the last to get in. Finally, reduced-form outcomes to school quality, such as aggregate productivity in the workforce or economic growth (Hanushek and Woessmann, 2012; Lucas, 1988) are beyond the scope of this paper.

My paper also contributes to the 'better schools' literature. Those studies (Clark, 2007; Jackson, 2010; Pop-Eleches and Urquiola, 2012) focus on attending better high schools and generally show that attending a higher quality high school leads to higher scores on a college entrance exam; a notable exception is Abdulkadiroğlu et al. (2011), which finds no benefit to attending elite exam schools in Boston and New York. An important different between Boston-New York study and the ones performed by Jackson as well as Pop-Eleches and Urquiola is that Abdulkadiroğlu et al. (2011) focuses on a very small subset of high-ability students at the very top high schools, whereas Jackson in Trinida-Tobago and Pop-Eleches and Urquiola in Romania look at the quality distribution of all students in the whole market of high schools in their market. Therefore, the latter two examples show that quality differences may be most important at lower points in the quality distribution. My paper on college quality shares this common focus on the full range of the quality distribution, instead of only looking at the top college(s) only (Saavedra, 2008; Hoekstra, 2009). Moreover, unlike papers in the 'better schools' literature, I can estimate wage returns in the labor market. And finally, unlike any other paper in the 'college quality' or 'better schools' literatures, I can do this using all college-goers in a cohort.

The remainder of the paper is organized as follows: section 1 describes the educational system in Taiwan,[2] the while section 2 details and summarizes the administrative data sets used. Section 3 motivates the use of RD estimation strategies. Section 4 discusses results, and Section 5 concludes.

2.1 Context

Taiwan's market for higher education is centrally controlled by the Ministry of Education (MOE), as is every level of education there. One of the historically distinctive characteristics of that centralized control has been its constraint of supply — both publicly and privately provided, and again at every level. This control extends down to the number of students and teachers that are allowed to be in each and every school, even at the academic department level, in any given year.

Early research on the returns to education in Taiwan has focused on the 1968 expansion of compulsory schooling to include the 3 years of junior high school. With this reform, junior high school enrollment increased by 50% in the first year (Clark and Hsieh, 2000). Several studies have looked at the effect of this junior high school expansion on such varied outcomes as the labor market returns to education (Spohr, 2003), the decrease in wage inequality (Vere, 2005), the increase in female labor market participation (Tsai et al., 2009), the effect of maternal education on infant health (Chou et al., 2010), and the intergenerational transfer of human capital (Tsai et al., 2011).

Besides the historically limited supply, the other distinctive characteristic of Taiwan's higher education market is constrained access, where admissions is determined by performance on a competitive examination and is centrally administered through a national, market-clearing student allocation system. Constrained access has received comparatively

[2] A more in-depth literature review on returns to quality is provided in Simpson (2012), or chapter 1 of this dissertation.

less attention in the empirical literature, mainly for lack of good microdata needed to facilitate such research. I attempt to fill this gap below.

Below is a description of the regular admissions process, which generates the discontinuities used in my research design. Both public and private institutions are subject to the centralized testing regime and the admissions protocol in Taiwan.[3] Students applying to college in the year 2000 have no other option for college enrollment than to follow Taiwan's prescribed protocol. Following this is a description of the matching algorithm used in the Taiwan case, including a hypothetical example to facilitate the discussion.

2.1.1 Regular Admissions

In the first week of July each year, any high school senior who hopes to attend college sits for the Joint College Entrance Exam (JCEE). The score on the college entrance exam T_i is a composite of subject-specific tests t_i. The number of subject tests a student must take ranges from 5 to 6 (max score for each is 100), and is predetermined by the MOE for each of the 4 educational tracks: social, science, medical, or agriculture.[4] (See Appendix Table A1.) Chinese and English are the only two subjects common to all the educational tracks. The scores on the JCEE T_i are a continuous measure, observed down to the hundredth of a decimal point with many data points occurring between any given test-point interval.

After observing her performance on the JCEE, the student must fill in an "aspirations card" that lists in ranked order her preferences for college departments.[5] In the notation that

[3]Admissions into a 2-year junior college or technical college is based on a separate, yet equally centralized admissions system.

[4]According to the MOE's system of categorizing college majors and academic departments that teach them, there are 4 major educational tracks. Within these, there are 8 fields of study, which can be further broken down into disciplines and subfields or college majors. Academic departments can then be defined narrowly by how the majors they offer are categorized (`www.edu.tw/files/site_content/B0013/bcode.xls`).

[5]Over the years the maximum number of preferences a student could list has increased. Currently, because

I use, I refrain from including an index for colleges because it simplifies notation. But more importantly, I do so because students cannot select a college only. The preference rank she lists must be for a specific academic department within a college. Thus, the discontinuities that I exploit are at the academic department level. Nevertheless, because of the research design I follow, I compare the quality of departments within a given college major; since for each college major a college typically has only one department, this means that the analysis reduces again down to a college quality comparison, taking into account college major. Because students are able to list preferences for academic departments in more than one college major, my results should be interpreted as a relevant estimate, based on one set of possible comparisons.

A student is only eligible for admissions at departments that are within the educational tracks she has chosen to pursue. Thus, the first rule of admissions is that a student must take the set of subject tests required for her desired college major. For example, if a student chooses to take the required subject tests for the science track, but not the tests for medicine, she will not be considered by the matching algorithm for any department of medicine. (More regarding the JCEE test and the admissions process is discussed in Section 2.4). This aspirations card is returned to the College Entrance Exam Center. The College Entrance Exam Center is the independent government agency attached to the MOE tasked with the administration of all elements of the college admissions process, from JCEE development to matching students to seats at university. (For ease of reading, hereafter I will use College Entrance Exam Center and MOE interchangably.)

---

students must list specific academic departments and not just colleges, the maximum is 80.

2.1.2 Admissions and the matching algorithm

The student matching algorithm (hereafter 'the algorithm') used by the MOE falls into a class of matching algorithms known as the Deferred Acceptance algorithm. Though the algorithm is used to prevent departments and students from choosing each other directly, it does allow for student preferences and department priorities to be taken into account during the matching process. To use the language of Abdulkadiroğlu and Sönmez (2003), departments have a set of "priorities" over student qualities they would like to enroll, whereas students have preferences for specific departments. Department priorities are more abstract in that they are not for specific students, but rather for qualities that any applicant may have. The matching algorithm uses the departments' priorities and students' preferences to match students to seats in specific departments.

The main objective of the algorithm is to allocate unoccupied seats in a specific department to students who have the highest priority and a preference for being in that department.[6] According to Abdulkadiroğlu and Sönmez (2003), the steps of the algorithm are as follows:

> "*Step 1*: Each student proposes to her first choice. Each [department] tentatively assigns its seats to its proposers one at a time following their priority order. Any remaining proposers are rejected.
>
> "In general, at
>
> "*Step k*: Each student who was rejected in the previous step proposes to her next choice. Each [department] considers the students it has been holding together with its new proposers and tentatively assigns its seats to these students one at a time following their priority order. Any remaining proposers are rejected" (p. 735).

[6] It is worth noting that the student is the initiator in the matching process, which gives him/her the first-mover advantage in creating a beneficial matching outcome.

This process terminates when all student preferences have been considered and all seats have been allocated.

The quote above suggests that departments are actively engaged in the priority ordering of students in each round. This is not truly the case in Taiwan. Departments notify the MOE of their priorities over student qualities before the admissions process begins and without ever seeing those who have applied to them. This can happen because in Taiwan's case, a student's priority order at a department is determined solely based on a student's JCEE scores. A department's priorities for specific student qualities are used by the algorithm to weight the exam performance of all students who have listed that department in their preference list; the weighting occurs in such a way that it influences the students' priority ordering for that department only.[7]

The outcome of the matching algorithm is final. Departments are required to admit all those that the algorithm assigns them; there are no wait-lists during the admissions process nor are there transfers between programs after admissions is over. The only way to be admitted into a department is to go through the MOE defined system. Students have no recourse if they do not like their match. Not liking their match does not mean that it was not their best possible match, given their performance on the JCEE and their priority ordering at departments listed in their preferences. If they so choose, they are allowed to retake the JCEE again the next year, but this is uncommon for admitted students.

According to Abdulkadiroğlu and Sönmez (2003) and many others, a Deferred Acceptance algorithm is a Pareto-efficient method for optimally matching one-to-one a finite number of students to a finite number of "objects" — in this case the objects are seats at college departments. Another advantage that this kind of matching algorithm has over other Pareto-efficient matching algorithms — such as the random serial dictatorship, which is often used

[7] This weighted exam score T_{id} which provides the priority ordering to the matching algorithm corresponds to the running variable in the RD design.

for such real-life purposes as allocating students to dorm rooms — is that the Deferred Acceptance algorithm allows for different schools (or college departments in the case of Taiwan) to have differing priorities over the same students. The same student may have different priority rankings at different schools, and the Deferred Acceptance algorithm — and thus my study — allows for these priorities to be accommodated.[8]

2.2 Data

This paper uses the universe of governmental administrative data on birth, education, and earnings outcomes for the 2000 cohort of college applicants in Taiwan. The micro-data are merged using national IDs. Table 1 provides the descriptive statistics. The sample of college admitted students appears to be evenly balanced in gender, and entering college at the expected age of around 18. Parents at the time of birth on average had less than a high school diploma, which is consistent with the historical expansion of public secondary education in Taiwan. Even though there has been worry that the market for college education has been expanding too quickly and thus lowering starting salaries are less than before, it appears that the first-year earnings of college graduates is still high, compared to their parents' income. The average first-year earnings for newly graduated students are only slightly lower than the average of their parents' combined income four years prior to the time the student took the college entrance exam. Thus, a college degree does appear to still have value in Taiwan.

[8]As Abdulkadiroğlu and Sönmez (2003) point out, a random lottery cannot take into account all the priorities that a department may have for students or priorities that are imposed upon them by an outside governing body. The example provides is of students being allocated to public schools, which often give higher priority to students who have siblings already in the school, who live within that school's catchment area, or who fall into certain mandated quotas.

2.2.1 Admissions data

The main data set is the college exam and admissions data from the 2000 application season (for students admitted to college for the 2000-01 academic year), which contains all records of exams taken. These detailed records include scores on each of the subject tests t that the applicant i took. The admissions data also includes the department ID and name (including the college's name) where the applicant was admitted, if admitted at all. To these department IDs and names, I merge other records from the MOE that contain the national versus private rank of the college as well as the tuition that was required for that department within that college. Approximately 111,150 students applied to college, of which 67,898 or 61% were admitted.

Measuring peer quality

The regular admission exams are made up of either 5 or 6 100-point subject-specific tests. The specific set of subject-tests a student is required to take is predetermined by the MOE for each educational track a student could pursue.[9] Thus, depending on the educational track a student pursues, students could earn a maximum of either 500 or 600 points. In order to use a common performance measure for students in the same college but different educational tracks, I use the percentage of total possible points scored:[10] this is calculated by dividing the individual scores by the total possible available (either 500 or 600, depending on what was required by the MOE) in that track and then multiplying by 100. From this, peer quality or ability (denoted as $\overline{A_i}$ in the formula below) can be calculated by averaging the percentage of total possible points scored at either college- or department-levels. The value

[9] See Table A1 in the Appendix for a detailed breakdown of which subject tests are required for each educational track.

[10] Note that the points described here are in terms of the UNWEIGHTED (or raw) test scores on the JCEE for each department, which is distinct from the WEIGHTED test scores used to make the running variable in the RD, described more fully below.

of n in the formula's denominator changes depending on whether the college or department level is used.

$$\overline{A_i} = \frac{\sum_{i=1}^{n}\left[\left(\frac{actual\,score_i}{total\,possible\,score_i}\right) * 100\right]}{n}$$

Identifying counterfactuals

The data does not supply the students' ranked preferences over all departments, only the name of the college-department where the student was admitted, if admitted at all. This realized outcome is only one of the several outcomes that we would like to observe in order to assess what could have happened. However, for the purposes of this paper, I make an assumption that students having been admitted into a particular academic department would have listed other academic departments in other colleges which also offered the same college major. Making this assumption allows information embedded in the examination and admissions records, as well as knowledge of the admissions process, to be used to generate a set of other departments offering the same college major, where the student would have been eligible to apply to depending on the educational track a student chooses: social, science, medicine or agriculture. Yet, even limiting the number of counterfactuals I consider to those departments where the student would have been eligible to apply to (based on the 5 or 6 subject tests she choose to take[11]) would still leave possibly hundreds of academic departments and dozens of college majors to consider for each student.

Pop-Eleches and Urquiola (2012) have a similarly unruly number of counterfactuals they could possibly consider since any prospective high school student in Romania is allowed to apply to any high school. Their solution is to limit the number of high schools (counterfactuals) they consider for each student to the ones in the village where the student lived.

[11] Few students choose the grueling option of taking all 9 subject tests to make them eligible for any department.

Following this guide to focus my analysis, I use the college major offered by the department where the student was admitted as a way to limit the number of counterfactuals I consider for that student. For example, for a student admitted into a department that teaches economics, I consider for that student only the potential outcomes in economics departments at other colleges. A limitation therefore of the data, and thus my study, is that a student's application strategy may be less oriented towards getting into a particular college major and more towards getting into a particular college. Therefore, my results should be interpreted as one of at least two possible and relevant counterfactuals to consider. However, because I eventually compare wage outcomes for these students, we expect that students graduating from the same major will have similar career paths and opportunities, and thus it is more preferable to compare within college major and across colleges rather than within colleges and across college majors.

2.2.2 Birth certificate data

I am able to match for 100% of the individuals in the admissions data to their birth certificate data, using the universe of all births in Taiwan for the birth cohorts 1978-1983. These data are collected by the Ministry of the Interior Affairs. From this data I get the student's date of birth (month and year), sex, and her parents' years of schooling (YOS) at the time of the child's birth.

2.2.3 Earnings data

The administrative data I use on earnings is for the years 2004 to 2005; 2004 is the first year those enrolled in college in 2000 would be able to find a full-time job. The data is administered by Taiwan's Bureau of Labor Insurance and collected by the Ministry of Labor Affairs. It is an employer-employee matched data set, which includes the employee's national

ID, monthly earnings for that year, and the employer's unique ID number.

As the name indicates, the Labor Insurance data (hereafter "the Labor data") is a record of healthcare insurance provided through employers. In 1995 Taiwan established universal healthcare coverage. All employed individuals were given access to this universal healthcare through their employer, and their monthly premia were paid half by the employer and half by the employee. The employee's share was based on a percentage of their monthly income. Thus, all employers were required by law to submit monthly earnings records for their employees in order for the premia to be removed from their earnings (Chou et al., 2003).

The data contain the complete universe of all private-sector workers in Taiwan, approximately 82% of the working population. Of those working adults not in the Labor data, about 10% are public sector employees, 3% are farmers/fishermen, and 5% are self-employed. The former three maintain a different form of public insurance that pre-date the 1995 universal healthcare. All unemployed are not in this data, because their unemployment benefits cover their healthcare provision (Chou et al., 2003).

A significant limitation of the Labor data to which I have access is that only about 33% of those in my sample show up in the earnings data for 2004 and 2005. This could be explained by several aspects of the average state of the labor market at this time. The main reason for the lower match rate may be due to a gender-specific binding obligations. Males who graduated from college in 2004 would likely have had to perform their 18-month national military service following graduation.[12] Since their military service would not end until December of 2005, we would not expect them to be looking for jobs until January 2006 at

[12]Reasons for not performing the service immediately (or at all) following graduation from college would be continued education at the post-graduate level, being deemed physically unfit for various reasons, being the only child (and thus would leave the parents without a child to assist them), having another family member currently servicing in the military, or fulfilling a bond with a government research position (`http://www.nca.gov.tw/ENGLISH/english.htm`).

the earliest. Thus, a desirable, though unfortunately infeasible, solution would be to get the Labor data for the 2006. For females, childbearing often follows immediately after marriage, since this is a common expectation in Asia. Thus, some females may not be represented in the labor data if they chose to stay at home.

A second reason for the lower match rate is that about 10% of all working adults hold a position in the government. This likely means that more than 10% of college graduates will work for the government, since the market for government jobs is consistently competitive requiring applicants to pass a very strict test, but rewarding those who succeed with higher than average wages, a stable career, and better benefits. Therefore, I assume that a large share of the government bureaucracy is staffed by college graduates. Lesser reasons could be that access to graduate school grew by about 3% a year, starting in 1999, and that the unemployment rate for college and university graduates rose to around 4% (Yang et al., 2010).

To check for evidence of whether this data limitation signals some element of selection, I regress an indicator for whether the individual has any wages on the model described further below. If we were to find a statistically significant (positive or negative) effect of having scored above the admissions threshold, then this would dampen our faith in the claims we might want to make about the outcomes of interest. If the effect were positive as well as statistically significant, then we would be concerned that any wage impact we observe would really be due to selection into finding a job. If it were negative, then we might say that there were positive selection out of the private sector into the public sector. The residuals of that regression are displayed graphically in the sixth panel of Figure 3. It shows that there is no statistically significant effect, meaning that the selection into employment and possibly even sector of employment is not driving our wage results. Even though the wage results do not appear to be driven by imbalance in the probability of public versus private employment, it does not diminish the fact that this smaller sample means that my reduced-form estimates are imprecise. Thus, if I find an effect, it should be more robust in the full data.

Because the earnings data is generated through the bureaucratic structures of the universal health-care system, it retains two key features of that system: it is measured at monthly intervals and somewhat roughly. The data I have access to is taken from the employee earnings records for the month of December for each year. There are advantages to this feature of the data. The first advantage of having monthly earnings from December is that we can capture earnings soon after graduates complete their education in June. The second advantage is that we record all observations on the same month of the year (in contrast to monthly earnings data recorded in the typical survey), which protects against the noise induced by business-cycle seasonality in monthly earnings. A possible disadvantage is that if a college graduate were between jobs in either December 2004 or December 2005, the data would not capture her employment. The only way to remedy this would be to have the monthly earnings for every month, which again is impossible to obtain.

Compared to other large national administrative records, the Labor data is measured fairly consistently (Figure 1). The upper limit in our records is 42,000 NTD, which is sufficiently high for most earnings of new graduates. Any wage above 42,000 will be top-coded at the upper limit. And the lower limit is 11,000 NTD, which again is well below the average earnings. Again, any wage below 11,000 will be bottom-coded as the lower limit. Lastly, the monthly earnings are listed in regular intervals. These characteristics of the Labor data — the truncated distribution and the intervals — are artifacts of the fee schedule for healthcare with a floor, cap, and menu-pricing scheme.

I use first-year earnings because it is a sensitive measure of employers' reliance on college quality for a rough signal of students' actual (but yet to be directly observed) productivity. If employers use the average quality of a student's college cohort to predict productivity, the first-year earnings should follow a similar pattern as the variance in these signals (MacLeod and Urquiola, 2011). Further, if a college graduate's starting salary is predictive of long-term earnings growth, starting on a lower (higher) rung of this pay scale likely means that longer term earnings will also be lower (higher) (Oreopoulos et al., 2012).

2.2.4 Data preparation

Let T_{id} represent the score of student i being considered by the matching algorithm for admissions to department d. T_{id} does not necessarily equal the simple summation of raw scores on all the subject tests that a student may have taken. More precisely, it is the combined realization of choices made by the student, the MOE, and the departments. Below I describe how each part of the following formula works to generate T_{id}:

$$T_{id} = \begin{cases} \sum_{i=1}^{N} \sum_{d=1}^{m} \sum_{j=1}^{J} t_i^j \rho_d^j w_d^j & \text{if} \quad t_i^j \geq \kappa_d^j \ \forall \ j \\ \oslash & \text{otherwise.} \end{cases}$$

The JCEE is a set J of 9 subject-specific tests created by the MOE, one subject test t^j for each subject $j \in J$. Of the 9 subject tests only 5 to 6 are required by the MOE for admission into one of 4 educational tracks: social, science, medicine, or agriculture. The MOE determines which subject tests are required for admission to any given department, according to the educational track the department falls into. Departments have no control over whether a subject j is required or not. Thus, the dummy ρ_d^j in the formula above indicates (0=no/1=yes) whether or not a specific j is required for that specific department $d \in m \in M$, which is a part of a specific college major m according to the MOE's set of college majors M.

Even though departments have no influence over required subject tests, the MOE does allow departments two ways to indirectly exercise their priorities for student quality: choosing a set of weights w_d^j and a set of subject-specific raw-score cutoffs κ_d^j to be applied to subject-specific test scores t_i^j from the JCEE. First, the MOE requires departments to choose weights w_d^j that increase the value of that required subject relative to other ones, allowing the department to emphasize that specific content knowledge over other content when summing them together. The menu of w consists of 1, 1.25, 1.5, 1.75, or 2. The default and most common weight is w=1, which leaves the raw score for that required subject score unchanged.[13]

Second, the MOE requires departments to choose from a menu of raw-score cutoffs κ_d^j to use on t_i^j. Raw-score cutoffs let departments truncate the distribution of viable candidates at some point in the distribution of t_i^j. Departments are not allowed to set where the cutoffs occur. The same four κ^j are available for any of the required subjects j. The default and most common raw-score cutoff is κ^j=0, which leaves the distribution of applicants unchanged. Specifically, the four κ^j are:

- A: only consider students with a t_i^j above the average of the upper 50th percentiles for subject j;
- B: only consider students with a t_i^j above the 50th percentile for subject j;
- C: only consider students with a t_i^j above the average of the lower 50th percentiles for subject j; or
- D: no raw score cutoff, which is analogous to κ^j=0.

Any student with a raw score t_i^j below the raw-score cutoff κ_d^j are automatically disqualified. In the data replication process I follow, if the condition $t_i^j \geq \kappa_d^j$ is not met, I approximate the algorithm's disqualification by giving i a missing value $\oslash$ for T_{id} for that department.

Notice that the formula above replicates counterfactuals only within college majors; for example, a student who is admitted into a linguistics department would have priority rankings replicated for her at other linguistics departments, but not a history department. In

[13]One could replace the ρ above with a w=0. But since departments have influence over which weights (w) to use but do not have influence over which subject tests are required (ρ), I choose to keep these features distinct in the model.

theory, I could replicate the counterfactuals for each student at every college-department d pairing within the specific college major m where they actually were enrolled. In practice, because I am considering the returns to attending a prestigious national college versus a private college within the same major, I replicate only one counterfactual per college major. It is the counterfactual associated with the cutoff for admissions into the academic department in the lowest ranked national college; that is, the cutoff is the lowest test score of any student admitted into a national college for that college major. I limit the counterfactuals in this way so that we can reasonably compare the monetary outcomes for those with the same college major.

To reconstruct T_{id} as it was generated by the student matching algorithm, I use the MOE's records of department priorities for 2000, which detail how each department chooses to use weights and raw-score cutoffs. The fact that the MOE allows each department to prioritize student quality along different margins of subject-specific knowledge means that there is a possibility that for each department on a student's preference list, she could have a different T_{id}. Still, even with the same score at any two departments, it is highly likely that the same student would have different priority orderings at those departments, if for no other reason than the two departments will likely not have the exact same mix of applicants. This corresponds to my description in Section 3 of how the matching algorithm generates for each student a unique priority ordering for each department in her preference list. I discuss below how this works in the context of multiple discontinuities being pooled together through a stacked RD framework.

Since the most common weights and raw-score cutoffs are the defaults, in many cases the total raw score and the total weighted score are the same. More importantly, even if a student's score were different for each of the departments she applied to, because the matching algorithm treats the ranking of students at each department independently, it only matters that within that specific counterfactual, all the students are being sorted along the same priority ordering (running variable), which uses the same required subject tests, weights, and raw-score cutoffs.

2.2.5 Matching example

To illustrate how required subject tests (ρ^j), weights (w^j), and raw-score cutoffs (κ^j) work in the student allocation process, take the hypothetical example with four students i_1, i_2, i_3, and i_4. Each is competing for the last seat in the same department. In this hypothetical example, the department's priorities for student quality lead it to value $t_i^{chinese}$ 1.5 times that of t_i^{math} and to want only students with a $t_i^{chinese}$ above $\kappa_d^{chinese(A)}$, that is, above the average of the top 50th percentiles in j=chinese.[14] Let us assume for the purposes of this hypothetical example that $\kappa_d^{chinese(A)}$=54 and that these two subjects are the only ones required for the matching to take place.

Looking at the matrix below, we can see that i_4 is automatically disqualified because he did not take the required math subject test. Next, we see that i_3 dominates in both raw and weighted total scores. However, because i_3's Chinese score is below the threshold $\kappa^{chinese(A)}$ for consideration, i_3 is also not admitted to d.

	$t_i^{chinese}$	t_i^{math}	Raw Total	Weighted Total	Enrolled?
i_1	64	66	131	162	yes
i_2	54	80	134	161	no
i_3	52	100	152	178$\Longrightarrow\oslash$	no
i_4	88	–	88	$\oslash$	no

Both i_1 and i_2 have $t_i^{chinese}$ above the raw-score threshold; and i_2 has the advantage in total raw score. However, because the department's priority weights Chinese by 1.5, i_1 narrowly beats i_2 for the seat. This is once again because i_1's total weighted score is higher, not

[14] This means that $w^{chinese}$=1.5, w^{math}=1, $\kappa^{chinese}$=A (or 54), and κ^{math}=D (or zero).

because i_1's $t_1^{chinese}$ is highest. Had i_2 been able to score 1.5 more points either by increasing her $t_2^{chinese}$ by 1 point or t_2^{math} by 2, she would have bested i_1 for the seat.

2.3 Empirical Strategy

Students are given the chance to select which departments they would like to attend. Because the number of seats in these departments is finite and predetermined under the MOE's supervision, this capacity constraint creates a cutoff c_d in access to department d at the JCEE score for the last student admitted. Such a selection rule makes this context amenable to using an RD design to estimate the value of gaining access to the department with the higher prestige (as well as higher average peer quality and better funding).

Since I do not have the full list of student preferences, I approximate this using a fuzzy RD design and interpret the resulting counterfactual as an intention-to-treat (ITT).[15] Specifically, Abdulkadiroğlu and Sönmez (2003) make a helpful distinction in their analysis of student matching algorithms. They show that a student may be eligible to be admitted to several schools according to the selection criteria created by each school; but this does not bind the student's choice. If a student is eligible for more than one school, she is given the one she prefers most, the one listed higher on her preference list. Thus, I make the distinction between being eligible for access and being admitted. A fuzzy RD takes into account that a student may be eligible for more than one, but may be admitted at only one.Below I motivate how this works with one discontinuity and then extend the framework to include many discontinuities.

[15] See Imbens and Lemieux (2008) for a technical guide to recent RD techniques and Lee and Lemieux (2010) for survey of recent uses in the literature.

2.3.1 Single discontinuity

Consider a hypothetical case in which there are two departments d_1 and d_2, ranked from lowest to highest by the quality (prestige and average peer quality) of the department. One cutoff c_1 defines the access threshold to d_2, the department with the higher quality. Students scoring above c_1 are eligible to attend d_1.

Let the following model (1) represent admissions outcomes — at both the first-stage and reduced-form stages — as a function of that student's achievement on the college entrance exam T_{id}:

$$Y_i = \beta 1(T_{id} \geq c_d) + b(T_{id} - c_d)^p + \gamma_c + X_i + \varepsilon_i \tag{2.1}$$

where $1(T_{id} \geq c_d)$ is an indicator function for whether a student's total weighted test score T_{id} is above the cutoff c_d for access to d. The function $b(T_{id} - c_d)^p$ is the summation of higher-order polynomials of order p, independently flexible above and below the cutscore. Within a narrow enough bandwidth around the cutoff (c_1 (in this hypothetical case), a flexible specification of $b(T_{id} - c_d)^p$ fully controls for other factors driving access to d_2. At wider bandwitdths in the running variable I use a 5th-order polynomial, then 3rd-order as I narrow the bandwidth, and at the narrowest bandwidths I use a linear fit. The error term ε_i is clustered at the department level, taking into account that students who applied to the same departments may have similar errors in the reporting of their record-keeping and/or some unobserved similarity.

A set of covariates X_i are included, but are not required for the coefficient of interest β to be causally identified. These include gender, age at time of taking the college entrance exam, and mother and father's years of schooling at time of birth. A main identifying assumption of RD is that individuals cannot manipulate their position on the running variable. The description of the rigors of the algorithm above should help to motivate this assumption theoretically. A practical test for this type of manipulation is to check whether or not student characteristics are balanced across the threshold. Even though it is fundamentally impossible to check whether all characteristics are balanced, I test for balance in the characteristics I have access to by inspecting whether there is overlap in their confidence intervals on either side of the discontinuity. Figure 3 displays the outcomes for the five covariates. The confidence intervals overlap, signifying that results are either statistically indistinguishable from zero. A sixth panel is provided in Figure 3 to test for whether we can predict whether someone ends up in my private sector wage data, given whether they are above or below the discontinuity. Again, there appears to be no differences between the characteristics on either side. The balance is the same for all five of the main covariates even when I condition the estimates of whether the sample had wages in my earnings data. From these results, we can infer that the covariates are balanced, and thus the positive effects we observe in the outcomes of interest are not due to substantive differences in student covariates around the cutoff.

In the first-stage analyses, we can interpret β as the expected improvement in college quality that a student experiences by scoring above the cutoff for access to d_2. As described above, I use three proxies for college quality. Because I am only considering the differentials in these three proxies at the discontinuity associated with admissions to (at least) the lowest quality national college for that college major, I interpret each of these outcomes as the benefits for admissions into a more prestigious college.

I use three proxies for college quality: the probability of being admitted into a national college $Pr(National)$, the average peer quality $\overline{A_i}$ at the college-level, and natural log of yearly tuition $ln(Tuition)$. Whether or not a student has a degree from a national college is an important label of success in Taiwan, as it is in other non-Western contexts (Rubinstein and Sekhri, 2011). Thus, I use $Pr(National)$ as an indicator for the college-prestige signal employers would observe when looking at an applicant's resume. A finer-grained measure of prestige (but one that is harder to observe with precision by the potential employer) is the average peer quality of the institution, proxied by $\overline{A_i}$; this measure does not necessarily distinguish between national versus private, but it estimates how individuals and cultures parse quality and allot prestige. Lastly, another valuable feature of attending a national college is that it is subsidized by the government, and thus would cost less than what private colleges offering the same major would cost (Liu et al., 2006).

In constructing the running variable for the RD, I use the weighted scores T_{id} because the algorithm uses the weighted score to sort students into priority ordering for seats, and thus the use of the weighted scores faithfully reconstructs the cutscore and the relative distance to the cutscore that each student would have experienced while the algorithm was allocating seats. However, the weights themselves have no substantive or economic meaning outside the allocation process, because it was the raw scores that students were given on the JCEE. Indeed, the weights are specific to the departments, and so it would be difficult to make comparisons of and interpret the difference in weighted scores across departments. Since subject-specific test scores t_i are the JCEE's measure of student performance, I use their unweighted total sum T_i^u (corresponding to the individual's actual score) to construct the average peer quality $\overline{A_i}$ for each academic department. The eventual outcome measure $\overline{A_i}$ has only a subscript for the individual i, because the only realization of average peer quality we are considering for each student is the one she actually experienced.

Taiwan's college admissions selection rule conditions access to any given department on the relative distance a student's test scores are above or below that department's ex-post realized cutoff. Such a rule is an as-good-as-random process. Thus, it allows my analysis the advantage of circumventing biases stemming from correlations a student's performance on the JCEE may have with other variables, which may also predict access to higher peer quality. For example, though we may be able to look at one specific student's performance on the JCEE and predict some likely level of college quality she would be eligible for, we could not WITH PRECISION predict her relative distance (or priority ordering) to the as-of-yet unrealized and unobserved cutoff for any given department. The realized distance of T_{id} from c_d (in either JCEE scores or rank scores) works as if it were arbitrarily set because it is a function of the ex-post realization of the cutoff c_d, which is due to factors beyond the control of any student or department. Such factors would include the number of students apply to that department and their test scores, that department's priorities for specific student qualities, and the capacity constraint set by the MOE for that department.

Now re-consider the reduced-form of model (1): where the outcome Y_i is again a function of a student's performance on the college entrance exam. The β coefficient on the term $1(T_{id} \geq c_1)$ now represents the increase in log-earnings (log-points) one is expected to receive from scoring just above c_1 compared to just below.

2.3.2 Multiple Discontinuities

My analysis exploits discontinuities for access into a national college for 35 college majors at 50 colleges, 21 of which are national colleges. Similar to Pop-Eleches and Urquiola (2012), I use a stacked RD set up. To my knowledge, aside from being the first paper to use a universe of applications within a contained college market, this paper is also the first analysis to use a stacked RD to estimate the monetary returns to college quality.

For each college major, scoring about the admissions cutoff to a national college is treated as a separate RD — a self-contained natural experiment with its own LATE estimate. By stacking all the RDs, I can aggregate all the LATEs to provide one estimate of the impact. Besides simplifying the task of interpreting the results, stacking RDs has the desirable characteristics of increasing statistical power by pooling observations. Stacking entails re-centering the running variable T_{id} for each student i. I do this by normalizing the student's test score the minimum threshold c_d required for admission to the academic department of lowest-ranked national college: $T_{id} - c_d$. I normalize test scores within college majors; thus, the subscript d in $T_{id} - c_d$ refers only to the academic department of lowest-ranked national college for that college major. When stacking RDs, it is common to include a fixed effect γ_c for each cutoff c_d to take into account that they occur at different points in the distribution of T.

Some college majors have no national colleges teaching it. Conversely, some college majors were only found in national colleges. Additionally, some private colleges are so far away in quality that when working even within a wider bandwidth, none of the students who were admitted into these private colleges fell within 80 points[16] of the admission cut off for the national college with the lowest rank[17]. In all three cases, it is impossible to compare the relative advantages of attending a national college versus not when there is no variation in national college status occurring within that college major or within the bandwidth of analysis. Thus, these college majors were explicitly dropped from the analysis when any disqualification applied. The result is that college majors in the following fields of study were completely dropped from the analysis: education, humanities and arts, and services.

Lastly, because re-centering and stacking the data multiple times aligns the discontinuities for each department at the cutoff, it increases the probability that there will always be an observation at T=0 and, therefore, mechanically increases the density of re-centered test scores at the cutscore. This artificial peak has nothing to do with students being able to manipulate which side of the cutoff they land on. Nevertheless, to allay any doubts, I follow the strict recommendations of Barreca et al. (2011), and I run a 'donut-RD' by dropping observations at T=0. Pop-Eleches and Urquiola (2012) also drop data at zero in their running variable.

[16] Almost 1.5 standard deviations of the running variable distribution of scores

[17] which proxied by admission cut off

2.4 Results

To summarize my results, scoring above the cutoff provides applicants with 50% greater chance of getting into a more prestigious college and have access to peers of 0.2 σ higher average ability. This quality improvement translates into a wage differential of around 0.7 log-points (on a 2-digit log base) or 0.22 σ in log earnings. All results are conditional on being admitted to college at all. They are robust across multiple specifications.

My preferred estimation strategy starts with a wider bandwidth on the running variable and works incrementally inward, cross-validating the results on a range of specifications (Imbens and Lemieux, 2008). The basic model fits a fifth-order polynomial in the running variable on the wider bandwidths. As the bandwidth narrows, I follow Angrist and Pischke (2009)'s recommendation to reduce the order of the polynomial, eventually ending up with a linear specification in order to improve the fit.

2.4.1 First-stage

Table 2 provides the first-stage results. The tabular results match well with what we see graphically in Figures 3-6, with each figure presenting residuals from different bandwidths and smoothing functions (linear or polynomial). Figures 3 and 4 present results using a linear fit and bandwidths of 20 and 30 points, respectively, around the cutoff. Figures 5 and 6 present results using a polynomial fit and bandwidths of 30 and 50 points, respectively, around the cutoff. All residuals are from a regression of the outcome variable (first-stage and reduced-form) onto a model including a linear trend in re-centered test scores, the covariates, and cutoff fixed effects.

The cross-validated results for impacts on prestige in Panel A (Columns 1-8 of Table 2 and Figures 3-6) look quite similar in size and significance across different bandwidths and specifications. These results in Panel A show that scoring just above the cutoff improves the probability of entering a national college by about 30-35 percentage points, depending on the specification, or increases the probability by about 50%. These are sizable differentials in prestige for a one-point difference in JCEE admissions score — which amounts to about 0.6 σ in prestige.

Panel B both in the Table 2 and Figures 3-6 show the impact on peer quality of scoring just above (versus below) the admissions cutoff for a national college in your college major. This specific outcome is more sensitive to functional form and bandwidth; this sensitivity to functional form can be seen most notably in Panel B of Figures 4 and 5, where switching from a linear specification in Figure 4 to a cubic one improves the fit immensely even on the same bandwidth of 30. This improved fit does not show up in the R-squared measures of the overall model, as seen in Columns 10 and 11 from Table 2; nevertheless, the improved fit is visually evident by comparing the graphical results. The improved model fit increases the magnitude and statistical significance of the discontinuity by more than 3 times to almost 2 percentage point increase in peer quality; the interpretation of the discontinuity is that for those scoring above the cutoff, one's classmates end up achieving 2 percent more of the total available scores on the college entrance exam. Though this may initially appear to be only a small improvement, when compared against the variance in the whole sample, it compares quite well. This 2 percentage-point increase is just under one-fifth of a standard deviation (0.2 σ).

Because the tuition a student pays each year is basically a function of whether or not she attended a national or private college, then it is unsurprising that Panel C looks to be the inverse of Panel A in each of the figures. Students who attend national colleges pay about half the amount that private students pay for the same college degree in the same college major (Liu et al., 2006). This is because the government heavily subsidizes national colleges. The discontinuity in the natural log of tuition is around -0.2 log points. That is, students who score just above the admissions cutoff for a national college are paying over 9000 NTDs less in tuition. Like its counterpart in probability of being admitted into a national college, the discontinuity in log tuition is equivalent to a 0.6 σ *decrease* in cost to students.

All three of these impacts suggest that admission into a national college carry with it different benefits experienced during the college-going years. These first-stage effects are likely to have several non-monetary and non-cognitive benefits for which I have no measure. For example, anecdotal evidence from around the world (but especially in East Asia) suggests that the prestige of the school where one's child attends is of major consumption value to parents. Thus, there is much pressure for children to succeed in this area. Being admitted into a prestigious college "gives face" to parents by vindicating their parenting skills. It also likely increases confidence in the student's own ability and future successes. In the next section I explore some evidence that these first-stage benefits also have reduced-form impacts on monetary returns.

2.4.2 Reduced-form

The first-stage impacts shown above correspond to sizable reduced-form impacts on log wages as well. Table 2 Panel D (Columns 33-40) shows that scoring just above the cutoff for admission into a national college for a college major improved log wages anywhere between 0.068 to 0.084, or about 0.22 σ in log wages. That is about 1000 to 1500 NTDs or 5% of average monthly earnings for this subpopulation of early graduates. Like the graphical results we saw for peer quality, the high-order polynomials provide a much better fit as the bandwidth increases. This can best be seen in Panel D as we move from Figure 4 to Figure 5, or from a linear fit to a cubic one. The improved fit increases both the size and significance of the coefficients as we move from Column 34 to Column 35.

Table 3 shows these results using different fixed effect and clustering specifications. Panel A is the same outcomes we saw in Panel D of Table 2. Panel B however changes the level at which the fixed effects occurs. Instead of using the fixed effect so that all those departments with the same cutoff share a common intercept (which assumes that they have roughly the same academic standards), letting each department have its own intercept highlights the major difference in national versus private college statuses. Across the board there is an increase in the impacts of about one-third to one-half larger sizes and slightly greater significance. Overall, these findings appear to be robust to specification testings.

2.5 Conclusion

Estimating the returns to a college education is a task common to both the general public and the research community alike. Families and governments around the world must decide how to allocate resources on behalf of their children and their citizens' education, with the hope that such educational investments will indeed payoff in the labor market and citizenry. As access to college has expanded around the world, it may not be enough to get just a college degree. To be competitive now, increasing numbers of students vie for seats at higher prestige institutions.

These thoughts motivate my estimation of the returns to college prestige. My paper's first contribution is to use data from Taiwan to overcome previous limitations to internal and external validity. As a context, Taiwan provides a highly centralized and very competitive market for students who are seeking admission to top schools, though not competitive for the universities itself. As a data source, Taiwan highly centralized government has been able to keep detailed and complete records of performance on college entrance exams, college admissions outcomes, and even earnings in labor market.[18] These factors and data provide a useful context in which college quality is finely sorted and its effects consistently measured.

[18] Thus, industrious researchers like Professor Jin-Tan Liu of National Taiwan University have been able to merge the data all together to great advantage.

This paper's second contribution is methodological. A stacked regression discontinuity design maps well onto the details of the student matching algorithm used in Taiwan, which sorts students by their performance on a central exam and matches their admissions preferences to seats in specific college departments. This research design allows me to efficiently summarize the impact of scoring above (below) an admission cutoff at any point in the entrance exam distribution, and then to relate that admissions outcome to the reduced-form impact in first-year earnings in the labor market.

Taken together, the analysis above presents results that are robust to a wide range of specifications and robustness checks. Despite the expansion of higher education, these positive and significant wage differentials indicate that there still is a college quality premium in the labor force. These positive wage returns to higher prestige would suggest that being the last admitted into an academic department in national college-department is (at least marginally) more preferable to being the first admitted into the next-best private college within the same major. In order for this to be generally true, it would suggest that employers in Taiwan are more sensitive to prestige signals than they are to other signals of college performance.

The prestige effects may not necessarily represent or be dependent on greater current learning occurring in national colleges. Upon entering the market, graduates from departments in more prestigious college offering the same college major could be benefiting from employer uncertainty regarding the graduates actual ability. If employers make make best possible guess using the average ability of a graduates cohort, even at the academic department level, then the marginally admitted student would be benefiting from the overall better signals of his/her more able peers. However, this does not have to be the case.

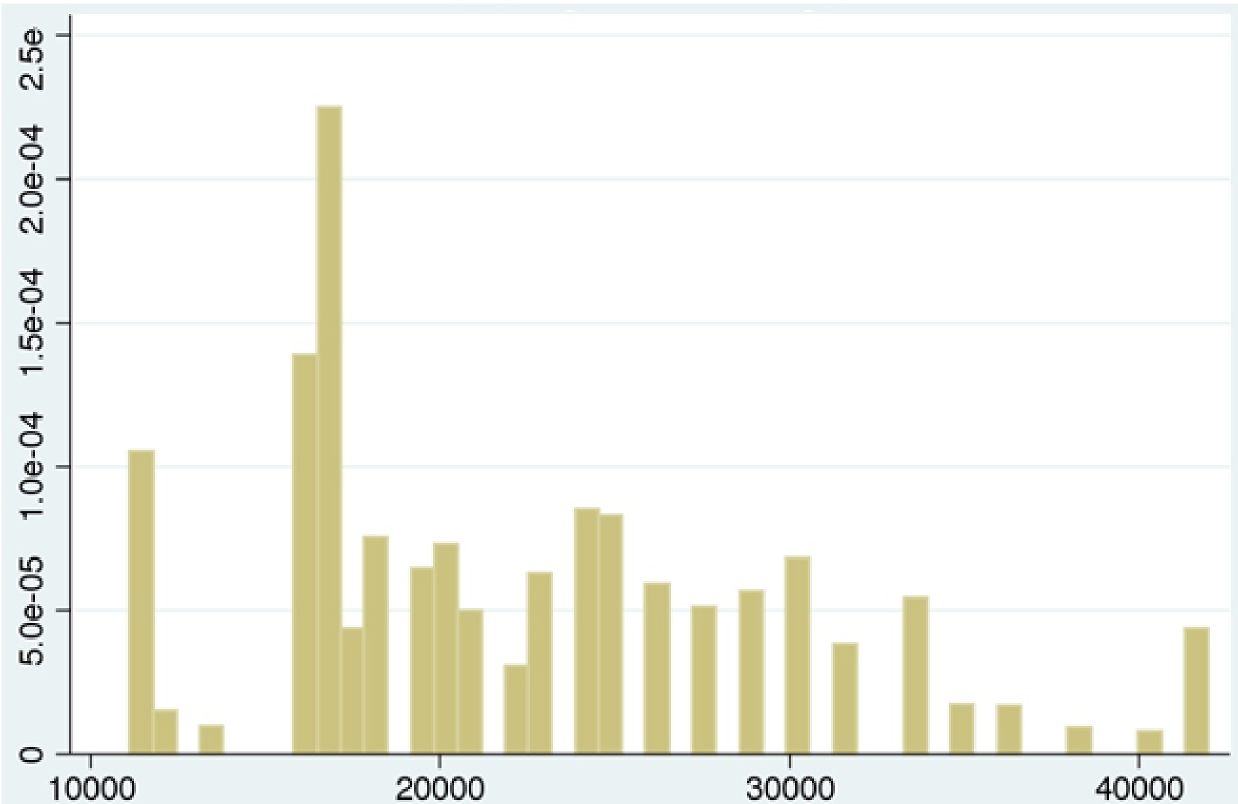
I suspect that the (jump) estimates I provide are lower bounds on the actual returns for two main reasons: Most importantly, I am able to merge only 33% of those in my education data with their earnings data. A larger portion of those missing from my earnings data are males, which earn more than females, according to previously literature of gender wage gaps in Taiwan (Tsai et al., 2009; Vere, 2005). If I were to have the Labor data for 2006, I believe that a significant share of these males would be showing up in my earnings data. Unfortunately, as I mentioned above, accessing the Labor data for 2006 at this point is unfeasible.

Increased first-year earnings is not the only benefit to higher college quality. The psychic benefits (to family) of attending a higher quality institution are important, especially when considering the reputation value of more prestigious ones. A full cost-benefit or welfare accounting would be useful, but are beyond the scope of this paper. Among other things, I would have to also account for the amount of money parents have spent on extra tutoring and/or post-high school cram schools as well as the foregone leisure students invested invested in studying in order to get into these higher quality schools. Given the intensity of the investments made by the family, a fair welfare analysis would also want to acknowledge the psychic benefits and costs. Since I do not have access to this information, I present these estimates of the savings on tuition to complement the wage differentials and as an acknowledgment of the overall welfare benefits.

The results in this paper may be specific to centralized systems like the one in Taiwan; nevertheless, such systems are common throughout the developing world where public higher education is rationed and thus very competitive. Moreover, these results provide a basis to further inspect the mechanisms that actually generate the positive relationship between college quality and earnings in such contexts.

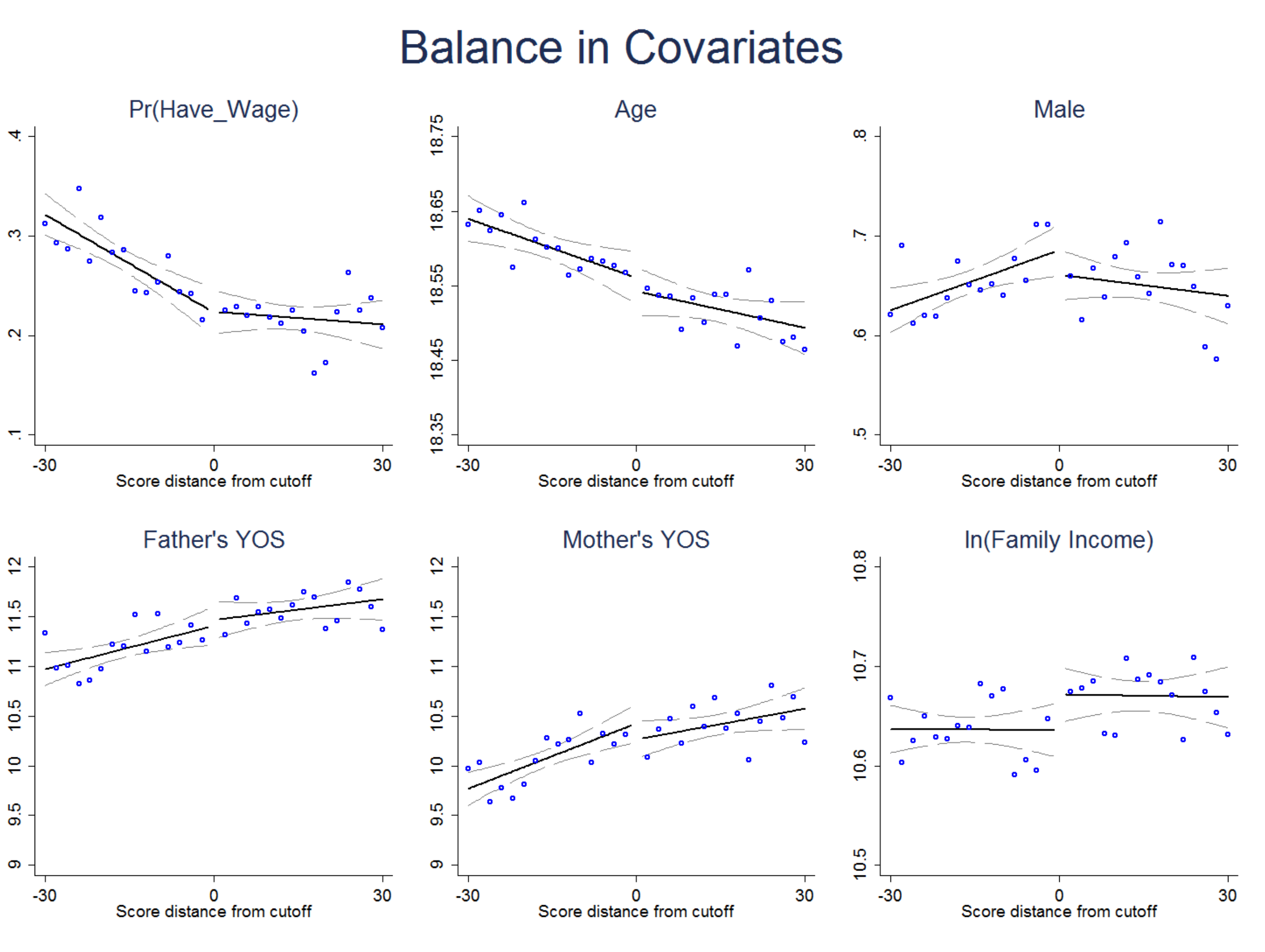
Figures

Figure 2.1: Histogram of wages



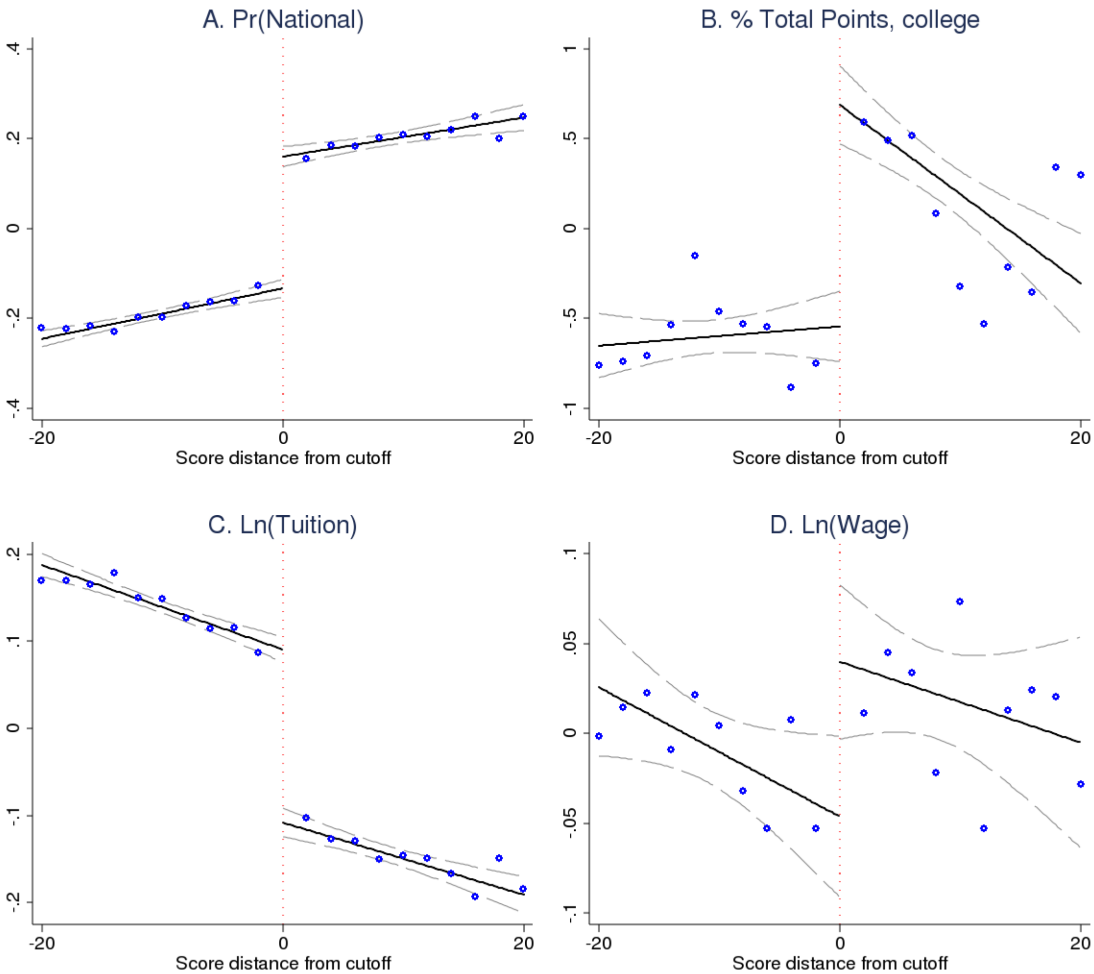
Notes: These represent the first year monthly earnings (in New Taiwan Dollars) for those entering college in 2000. The data were recorded in the month of December for 2004 and 2005. The data are from the Labor Insurance scheme that Taiwan follows in which an individual's premium is a function of his/her monthly salary. Thus, because of the intervals in the menu pricing, the wage records are also lumpy and truncated to a range between 11,000 NTD to 42,000 NTD.

Figure 2.2: Balance in the Covariates



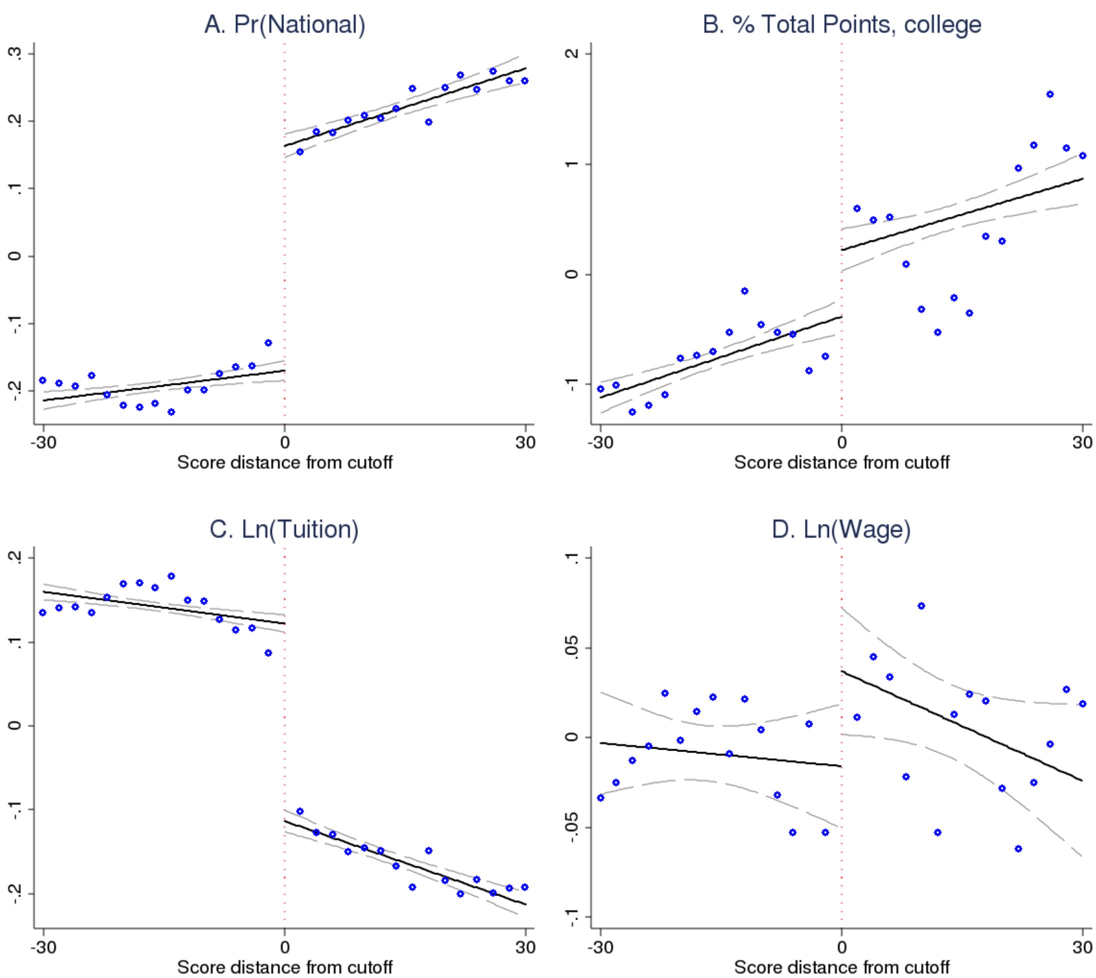
Notes: The plots above present the linearly smoothed fits of each control variable around the admissions cutoff. The dots represent the bin averages using a 2-point bin-width.

Figure 2.3: Basic results with bandwidth of 20 and linear fit



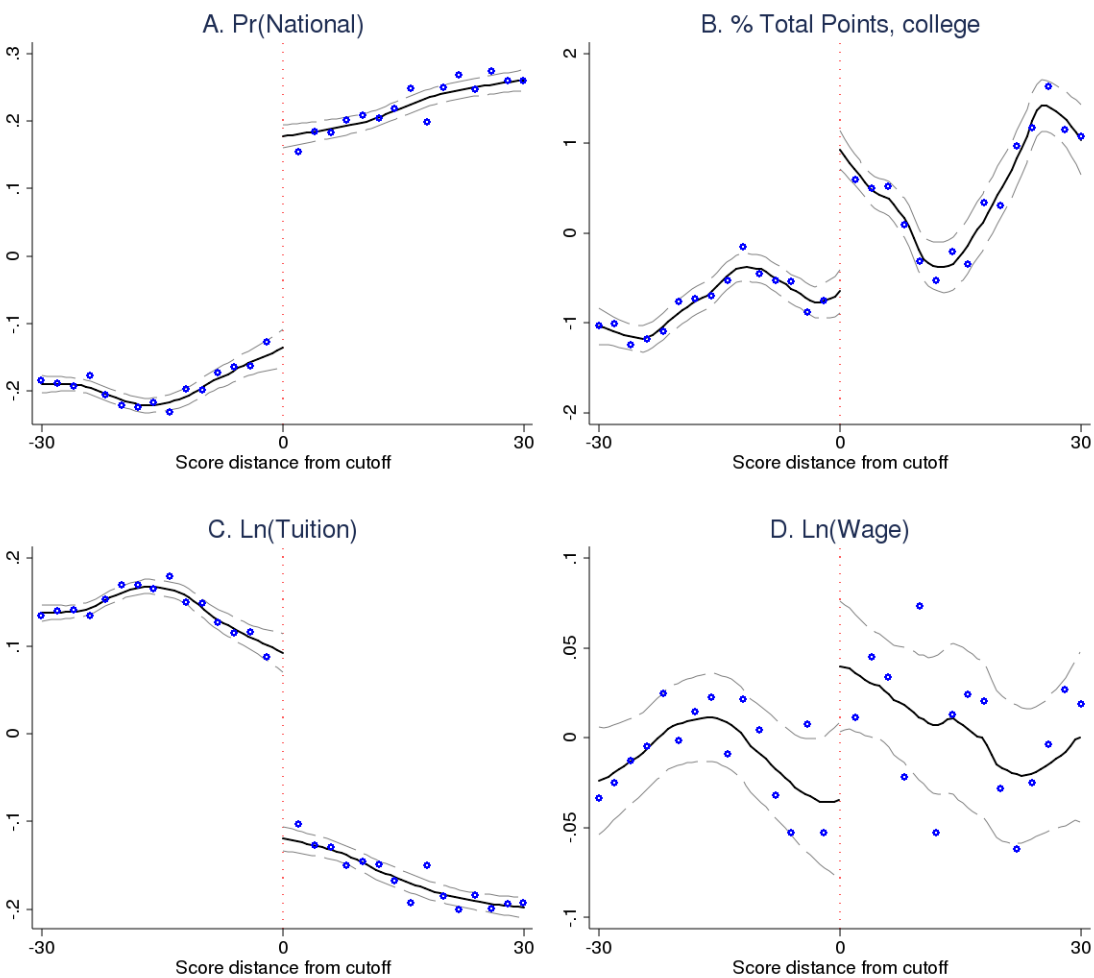
Notes: The plots above present the linearly smoothed fits of residuals from a regression of each outcome variable (first-stage and reduced-form) on to a model including a linear trend in re-centered test scores, the covariates, and cutoff fixed effects. The dots represent the bin averages of these residuals using a 2-point bin-width. Discontinuities correspond to beta-coefficients in Columns 1, 9, 25, and 33 of Table 2. First-stage results are the first 3 panels: "Pr(National)", "% Total Possible @ College", and "Ln(Tuition)". The reduced-form result is the last panel: "Ln(Wage)."

Figure 2.4: Basic results with bandwidth of 30 and linear fit



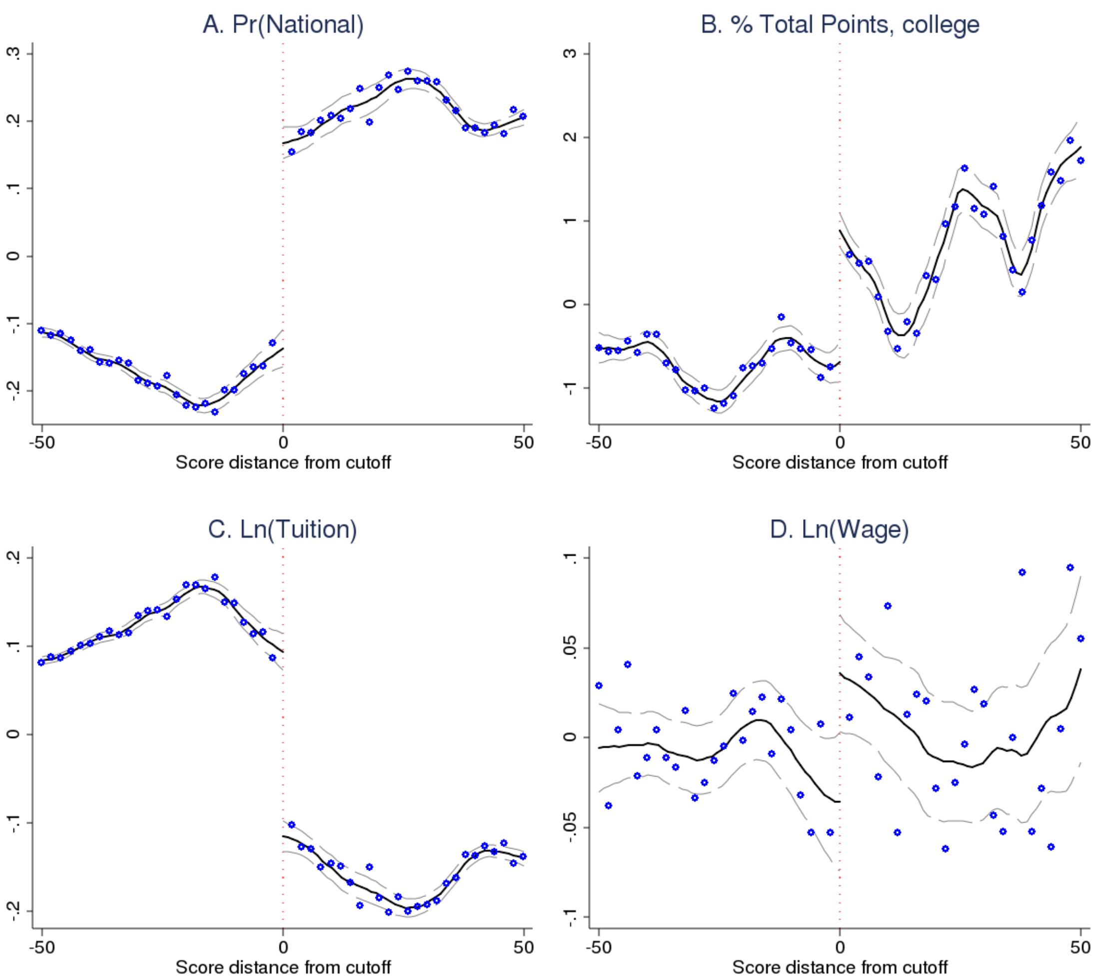
Notes: The plots above present the linearly smoothed fits of residuals from a regression of each outcome variable (first-stage and reduced-form) on to a model including a linear trend in re-centered test scores, the covariates, and cutoff fixed effects. The dots represent the bin averages of these residuals using a 2-point bin-width. Discontinuities correspond to beta-coefficients in Columns 2, 10, 26, and 34 of Table 2. First-stage results are the first 3 panels: "Pr(National)", "% Total Possible @ College", and "Ln(Tuition)". The reduced-form result is the last panel: "Ln(Wage)."

Figure 2.5: Basic results with bandwidth of 30 and polynomial fit



Notes: The plots above present the polynomial smoothed fits of residuals from a regression of each outcome variable (first-stage and reduced-form) on to a model including a linear trend in re-centered test scores, the covariates, and cutoff fixed effects. The dots represent the bin averages of these residuals using a 2-point bin-width. Discontinuities correspond to beta-coefficients in Columns 3, 11, 27, and 35 of Table 2. First-stage results are the first 3 panels: "Pr(National)", "% Total Possible @ College", and "Ln(Tuition)". The reduced-form result is the last panel: "Ln(Wage)."

Figure 2.6: Basic results with bandwidth of 50 and polynomial fit



Notes: The plots above present the polynomial smoothed fits of residuals from a regression of each outcome variable (first-stage and reduced-form) on to a model including a linear trend in re-centered test scores, the covariates, and cutoff fixed effects. The dots represent the bin averages of these residuals using a 2-point bin-width. Discontinuities correspond to beta-coefficients in Columns 5, 13, 29, and 37 of Table 2. First-stage results are the first 3 panels: "Pr(National)", "% Total Possible @ College", and "Ln(Tuition)". The reduced-form result is the last panel: "Ln(Wage)."

Tables

Table 2.1: Descriptive Statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
A. Outcomes					
Log Wage	11255	10.01	0.35	9.32	10.65
Peer Quality @ College	39951	51.63	9.45	34.57	71.28
Pr(National)	39951	0.29	0.45	0	1
Log Tuition	39951	10.57	0.32	9.89	11.03
B. Covariates					
Gender	39951	0.64	0.48	1	
Age	39951	18.64	0.67	15.67	22.50
Mother's YOS	39925	11.23	3.50	0	18
Father's YOS	39917	10.06	3.54	0	18

Table 2.2: Cross-Validated Results: First Stage and Reduced Form

A. Pr(National)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
1(Treat)	0.247***	0.305***	0.311***	0.335***	0.329***	0.329***	0.354***	0.371***
	(0.065)	(0.065)	(0.066)	(0.067)	(0.069)	(0.071)	(0.071)	(0.071)
bwidth	20	30	30	40	50	60	70	80
poly	1	1	3	3	5	5	5	5
R2	0.59	0.63	0.63	0.67	0.71	0.74	0.76	0.77
N	7689	11620	11620	15904	20181	24040	27765	31156

B. % Total Possible @ College

	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
1(Treat)	1.049*	0.624	1.977***	1.592***	2.423***	2.093***	2.240***	2.091***
	(0.497)	(0.520)	(0.535)	(0.550)	(0.592)	(0.612)	(0.625)	(0.628)
bwidth	20	30	30	40	50	60	70	80
poly	1	1	3	3	5	5	5	5
R2	0.73	0.73	0.74	0.77	0.80	0.81	0.83	0.84
N	7689	11620	11620	15904	20181	24040	27765	31156

C. Log Tuition

	(25)	(26)	(27)	(28)	(29)	(30)	(31)	(32)
1(Treat)	-0.165***	-0.217***	-0.212***	-0.233***	-0.222***	-0.225***	-0.242***	-0.255***
	(0.047)	(0.048)	(0.048)	(0.049)	(0.050)	(0.051)	(0.052)	(0.052)
bwidth	20	30	30	40	50	60	70	80
poly	1	1	3	3	5	5	5	5
R2	0.60	0.64	0.64	0.67	0.71	0.73	0.75	0.76
N	7689	11620	11620	15904	20181	24040	27765	31156

D. Log Wage

	(33)	(34)	(35)	(36)	(37)	(38)	(39)	(40)
1(Treat)	0.068*	0.050+	0.075*	0.066*	0.084*	0.064+	0.071*	0.069*
	(0.031)	(0.026)	(0.031)	(0.029)	(0.042)	(0.038)	(0.036)	(0.034)
bwidth	20	30	30	40	50	60	70	80
poly	1	1	3	3	5	5	5	5
R2	0.23	0.22	0.22	0.21	0.19	0.18	0.18	0.18
N	1838	2917	2917	4105	5307	6513	7722	8817

Notes: The tabular results above present the results from regressions each outcome variable (first-stage and reduced-form) on a model including an indicator for being above or below the admission cutoff, and function of the running variable, covariates, and cutoff fixed effects. Error term was clustered at the department level. Panel A represents prestige results for the probability of being admitted into a national college within your college majors for scoring at or above the admissions cutoff for the national college with the lowest admissions cutoff in that college major. Panel B is for peer quality results at the college-level. Panel C is for institutional benefits in the form of reduced tuition. Panel D is the reduced-form benefit in the form of first-year log wages. The results are cross-validated on various bandwidths ("bwidth") and polynomial fits ("poly"). Standard errors in parentheses. + $p<0.10$, * $p<0.05$, *** $p<0.01$.

Table 2.3: Cross-Validated Results with Different Fixed Effects and Clustering Specifications

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Panel A: Cut-off Fixed Effects							
1(Treat)	0.068*	0.050+	0.075*	0.066*	0.084*	0.064+	0.071*	0.069*
	(0.031)	(0.026)	(0.031)	(0.029)	(0.042)	(0.038)	(0.036)	(0.034)
bwidth	20	30	30	40	50	60	70	80
poly	1	1	3	3	5	5	5	5
FE	Cut	Cut	Cut	Cut	Cut	Cut	Cut	Cut
R2	0.23	0.22	0.22	0.21	0.19	0.18	0.18	0.18
N	1838	2917	2917	4105	5307	6513	7722	8817
Clusters	213	278	278	335	392	440	483	513
	Panel B: Department Fixed Effects							
	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
1(Treat)	0.083*	0.064+	0.101***	0.093*	0.116*	0.094*	0.099*	0.105***
	(0.041)	(0.037)	(0.036)	(0.035)	(0.047)	(0.037)	(0.036)	(0.037)
bwidth	20	30	30	40	50	60	70	80
poly	1	1	3	3	5	5	5	5
FE	Dept	Dept	Dept	Dept	Dept	Dept	Dept	Dept
R2	0.34	0.31	0.31	0.30	0.28	0.27	0.26	0.26
N	1838	2917	2917	4105	5307	6513	7722	8817
Clusters	35	35	35	35	35	35	35	35

Notes: The tabular results above present the results from regressions each outcome variable (first-stage and reduced-form) on a model including an indicator for being above or below the admission cutoff, and function of the running variable, covariates, and fixed effects. The results are cross-validated on various bandwidths ("bwidth") and polynomial fits ("poly") and fixed effects ("FE"). Panel A uses cutoff fixed effects and clustered standard errors at the department level as those presented in Table 2. Panel B uses department level fixed effects and clustered standard errors at the college major level. Standard errors in parentheses. + $p<0.10$, * $p<0.05$, *** $p<0.01$

Chapter 3

A College By Any Other Name: Returns, Reputation, and Reform

"What's in a name? That which we call a rose
By any other name would smell as sweet." –*Romeo and Juliet*

3.1 Introduction

Despite an increase in access, competition for quality in higher education has continued to increase substantially over the last few decades. Institutions known for being selective are now even more selective, measured in terms of both the increase in applications to selective schools (Bound et al., 2009) as well as the resulting increase in rejection rates (Vigdor and Clotfelter, 2003). Why has there been this increased preference for college reputation?

One reason is that as more individuals continue on for further education, educational attainment becomes less of a distinction. Students can, of course, press on for even more years of schooling, but there may be diminishing marginal returns to this strategy. Thus, switching from an external margin of competition (through quantity) to an internal margin of competition (through quality) may be motivated by the belief that returns to higher quality schools persist even when the overall supply of graduates increases.

In addition to the possibility that returns to quality are higher and/or persistent, there are other ways in which school reputation could be valuable in the market place. One way is that a college's reputation helps resolve the information problem that students and employers face in the hiring process (MacLeod and Urquiola, 2011). Students need some way to evidence their potential and employers need some way to observe it. Using a school's reputation, prospective students (and their parents) have a way to estimate how worthwhile their attendance at a particular school would be, on average, based on the past performance of previous graduates from that institution; likewise, prospective employers have a way to estimate how productive applicants from a particular school would be, on average, based again on past experience. In this way, a college over time builds a reputation based on the expected returns its graduates achieve in a competitive labor market. Thus, if a school's reputation represents the quality of its student body, then graduates from schools with higher reputations should earn higher returns.

Reputation could have this ameliorating effect on the hiring process (and wages) even if there is no evidence that the school has a positive value-add on learning. In other words, education (and schools) need not cause students to be more productive employees later on; education could simply predict it (Chevalier and Conlon, 2003). Suppose the school does not increase the student's stock of human capital, it still can reduce employers' effort in identifying employment candidates, while improving the credibility of the student's own signal of her employability (Spence, 1973; Weiss, 1995). Conversely, if schools can increase students' human capital, then a school's reputation imparts this information as well. Arcidiacono et al. (2010) argue for this latter interpretation of education as both a skill-augmenter and a signal-refiner. In like manner, (MacLeod and Urquiola, 2011) assume in their model of school reputation both human capital and signaling are involved.

Much of the previous literature has measured quality or reputation in terms of selectivity. Selectivity is relevant in western contexts where a strong public university system satisfies the demand for higher education. But in many developing countries, publicly provided education is rationed, leaving the (often lower-quality) private market to satisfy excess demand (Rubinstein and Sekhri, 2011). Indeed, in my research context of Taiwan, all access to post-secondary education has been rationed, for both public and private institutions. Consequently, measuring school reputation in Taiwan in terms of selective versus non-selective institutions is uninformative, because the whole higher education market in Taiwan is selective.

Therefore, this work focuses on the reputational distinction between institutions that are of college-rank versus university-rank. The labels ‘college’ and ‘university’ are important reputation signals for two reasons: first, it is one of the most immediately understood and immediately available pieces of data an employer has for making hiring decisions. It is the most immediately available because it is one of the primary components of a resumé for a college graduate. It is the most immediately understood because the hierarchical distinction between colleges and universities is a simple notion. The labels ‘college’ and ‘university’ usually represent the fact that a particular higher education institution (referred to as “institution” hereafter to maintain the distinctiveness of the labels ‘college’ and ‘university’) has met some minimum standard of quality, where a college must meet a lower minimum standard than a university does. In my context, the highest degree a college can confer is centrally capped at Bachelor degrees, whereas universities can confer postgraduate degrees. Therefore, if employers are sensitive to these signals, a college degree should be of lesser value than a university degree.

To preview the results, the upgrading of colleges to universities increased the probability of attending any higher education institution by 4.5% and the probability of attending a university instead of college by (a very conservative) 9.5% (or Wald estimate of 20.45%). This translates into a 3 to 6 percentage point advantage to first-year earnings of attending a university compared to a college. These results are robust to multiple analytical techniques and specifications.

The rest of the paper is as follows: Section 2 lays out the literatures on school quality/reputation and the returns to education in Taiwan. Section 3 describes the data; Section 4 deals with identification; Section 5 presents and discusses the results. The paper finishes with concluding remarks.

3.2 Literature Review

I review the literature on school quality, highlighting its range of human capital explanations for the effects of education on labor market performance. I also summarize previous work on the monetary and non-monetary returns to education in Taiwan.

3.2.1 School Quality

The recent 'better school' literature has produced several noteworthy papers, focusing on better peer quality as the mechanism for human capital benefit. In addition to being recent, its relative importance to this study is that it is the main outlet for papers on higher education.[1] These papers seek to answer the question: what are the human capital returns to attending a better school/college? None has tackled questions regarding the signaling value of school reputation.

The 'better schools' literature has offered compelling research designs to circumvent previously insurmountable forms of selection, thus identifying credible, causal estimates of the human capital effects of school quality on later-life outcomes. One of the earliest example of the school quality literature is Dale and Krueger (2002). They exploit a unique dataset of students who applied to similarly selective universities: some applicants eventually attended the more selective school, while others chose to attend the less selective ones. With this data, Dale and Krueger attempt to control for two critical and often ignored types of selection: (1) selection on the part of the university in choosing enrollees from all their applicants, and (2) the subsequent selection by enrollees (and their parents) on which institution to attend. They find that those who attended the more selective institution enjoyed an earnings advantage upon graduation, but this advantage wore off after about 5 years, implying that selective institutions do not impart any lasting advantage, once selection has

[1] Previous work on resources and teacher quality papers have focused exclusively on basic education.

been appropriately controlled for.

MacLeod and Urquiola (2011) interpret these results, however, as suggesting that employers were using the college's reputation as an initial measure of the graduate's quality. Upon hiring the graduates, though, employers were able to directly observe an individual's productivity and could adjust the wage accordingly.

Another set of contributions in the 'better schools' literature implement regression-discontinuity (RD) designs to infer causation.[2] The papers listed below can be categorized into (1) those RD designs estimating the value of going to a better high school, and (2) those estimating the returns to going to a better college. We take each in turn, starting with the return to going to a better high school.

Jackson (2010) shows that in Trinidad and Tobago, students scoring just high enough to be admitted into a selective school experienced higher exam scores on the college entrance exam, relative to those who did not score high enough and attended a less selective high school. The quality of data in this analysis is exceptional since it includes the full listing of all schools the applicant applied to, in order of preference. No other paper has this level of detailed information. It is an extreme version of the control strategy that Dale and Krueger attempted, because Jackson has the universe of all test-takers and the complete set of all their desired schools. Therefore, his results offer compelling evidence of the effect of attending a higher quality school, as measured by the average entrance exam score of one's peers.

Clark (2007) looks at the public high school context in the UK. He finds that getting into a high-achieving school by scoring just above the cut-off on an IQ test in the last year of primary school does not improve one's standardized test scores at the end of high school. However, it does increase the likelihood of taking more difficult advanced placement courses and eventually attending college. This is exactly the opposite finding from Jackson (2010).

[2]For a methodological summary of regression discontinuity designs, see Imbens and Lemieux (2008); and for a review of how RDs have been used in economics, see Lee and Lemieux (2010).

Finally, using an RD design with over 2000 discontinuities across the full distribution of student achievement, Pop-Eleches and Urquiola (2012) show that those who just make it into a higher quality high school, relative to those who just miss it, have approximately 0.1 standard deviations higher average peer quality. The higher average peer quality translates into 0.02 to 0.10 standard deviations higher scores on college entrance exams, depending on whether one looks at the administrative or survey data. These results encourage the interpretation that schools do have a small, yet highly significant effect.[3] Based on data from an extensive follow-up survey, Pop-Eleches and Urquiola show that students who just make it into a higher achieving school experience lower self-esteem in the early years of the transition, though eventually recover. They also study how parents and schools adjust. For example, when a student makes it into a higher achieving school, her parents also reduce effort in helping her study. Conversely, higher achieving schools experienced increased teacher quality, because higher-qualified teachers prefer to teach higher-achieving students. These two opposing forces of reduced parental effort and increased teacher quality may balance each other out. We return to these behavioral outcomes later on in relation to our identification strategy.

There are currently only three papers that estimate the returns to going to a better college; only two of them contain earnings data. The earliest example with earnings is Hoekstra (2009), which shows that attending a flagship university in Texas pays approximately 18% higher returns up to 10 years after graduation, compared to those who just missed the admissions cutoff for that flagship college. This finding is a reversal of Dale and Krueger's findings, and is arguably a more credible estimate for its use of the more rigorous RD research design.

In addition to being the first to complete the difficult task of linking college admissions data to later earnings data at the individual level, another major contribution of Hoekstra

[3] The small size of the estimate can to some degree be explained by the fine-grained nature of the differentiation between schools.

(2009) lies in the quality of the earnings data Hoekstra was able to obtain. Most papers use only survey data, while Hoekstra uses administrative records. Second, most papers use only one observation on earnings, commonly being first-year earnings. Using first-year earnings is not necessarily misleading or insufficient, but it would be informative to learn about education's long-term effects. Hoekstra's data are exceptional in two ways: First, his data include earnings ten years after graduation, so that the estimates give more of a lifetime earnings interpretation. Second, by getting quarterly income tax records for a period of 5 quarters, the estimate is not as erratic as a single measure of a single point in time. The paper is not without concerns, however. In his RD estimation strategy, Hoekstra does not know where those just below the cut-off for admission into the flagship university eventually went to college, or even if they did. This makes the generalizability of Hoekstra's results especially limited. Nevertheless, as the first RD with earnings and the advantages mentioned above, it represents an important reassessment of Dale and Krueger's findings on the value of college selectivity.

Saavedra (2008) is a strong second example of an RD on the academic and labor market effects of attending a higher quality university. His natural experiment provides a rare opportunity to capture value-added at the tertiary level because the Colombian university system employs both an entrance and a subject-specific exit exam. Saavedra uses two strategies. First, he uses the detailed list of all applicants for one of the top ranked universities in Colombia (Los Andes University). Those who score just above the admissions cut-off enjoy about a 0.5 standard deviation improvement in peer quality (and receive 40% higher per-pupil expenditures) relative to those who were just below the admissions cut-off. These benefits translate into a 0.2 standard deviation improvement on subject-specific *exit* scores. Subsequent to graduation, those just above the cutoff are 16% more likely to be employed and earn 35% more.[4] Second, Saavedra infers the admissions cut-off for the 25 most selective institutions in Colombia, and he finds qualitatively similar positive results. These are all statistically and substantially significant results.

While the RD examples above show a positive effect of attending a selective school, Rubinstein and Sekhri (2011) find little evidence for a positive effect on educational outcomes. Like Saavadra's Colombian context, higher education in India has both entrance and exit exams. Using this feature in a similar RD design, Rubinstein and Sekhri (2011) show that just making it into public universities, which are more selective than private ones, provides no improvement on subject-specific exit exam scores, compared to those who were just below and had to attend private schools. Unfortunately, since they do not have any earnings data in their analysis, we are unable to see if there are other non-academic employment or income advantages to attending a more selective (though arguably similar in value-added) institution.

This paper is also related to the literatures on peer-effects, which focus on measuring the extent to which achievement gains can be attributed to a student's classmates. The major objective of the peer quality literature, beyond establishing the existence of peer effects, is to test for nonlinearities or multiple equilibria in peer quality. If peer quality is a linear function of student characteristics, then the policies that promote tracking are simply zero-sum games (sometimes called cream-skimming) in which one student must be made worse off to make another student better off. Two randomized experiments provide opposing views on the policy implications of this question. First, Duflo et al. (2011) show that in Kenya tracking students by ability into two groups allowed each group to improve their overall achievement relative to the control groups that were not tracked. They argue that tracking helped students receive more targeted teaching because there was less variance around their

[4] Saavedra points out that once he conditions the wage regression on being employed, there are no wage differentials to attending the better college, which seems to show that the main effect was to employment.

average levels. Carrell et al. (2011) caution, though, that even well-informed, well-meaning experts cannot foresee all of the effects of intervening in peer dynamics or forestall unintended consequences.

3.2.2 Returns to education in Taiwan

Education at all levels in Taiwan has historically been limited in supply, and thus very selective. For example, during Taiwan's pre-industrial period, public education was provided only through primary school. Starting in 1968, there was a massive increase of public education through the numeric and spatial expansion of junior high schools and the elimination of the previously mandatory junior high school entrance exam. Consequently, the junior high school enrollment increased by 50% in the first year (Clark and Hsieh, 2000). Several studies have looked at the effect of this junior high school expansion on such varied outcomes as the labor market returns to education (Spohr, 2003), the decrease in wage inequality (Vere, 2005), the increase in female labor participation (Tsai et al., 2009), the effect of maternal education on infant health (Chou et al., 2010), and the intergenerational transfer of human capital (Tsai et al., 2011).

There has been a similarly large excess demand for higher education. Two dynamics are evident from looking at Figure 1: (1) the higher education school system after 1975 in Taiwan was quite stable into the early 1990s, and (2) many existing institutions were allowed to upgrade to college- or university-status over periods 1997-2006. The government's policy throughout the post-1971 industrial period was to channel public interest in post-compulsory education toward the acquisition of vocational skills. As part of this plan, the government allowed a rather significant increase in the number of junior colleges, especially private ones, from just under 20 in 1966 to over 70 in 1971. It was a very constrained expansion to meet centrally devised policy targets: namely, a 60:40 vocational to general tracks ratio in 1972, which was updated to 70:30 in 1980. Consequently, during this time, the number of colleges and university (that is, general education) remained relatively unchanged (Tsai and Shavit, 2007).

Figure 1 illustrates how following this 'big push' in vocational training, the supply of higher education has remained stabilized at its 1975 levels by the central government until the early 1990s, with only slight adjustment after the demilitarization of Taiwan in 1987. Gindling and Sun (2002) use this slight increase in the number of institutions following 1987 to estimate the change in returns to higher education; they find a negative (sometimes significant) relationship.

3.3 Data

The data sets used in this paper represent the administrative birth, education, military , and earnings records for the universe of all surviving children from the 1978 to 1983 birth cohorts. The different data sets are merged using the unique national identification number of each individual. The totals number of observations is 2,379,470 individuals. Few data sets exist for academic research purposes that match so many important life events for so many individuals.

3.3.1 Birth Data

The data, collected by the Ministry of the Interior Affairs, comes from the annual birth certificate records for 1978 to 1982. Though more variables are included in the birth certificate data, I consider the following variables in my analysis: gender, year- and month-of-birth, county-of-residence, urban status, parity, birth weight, gestational age in weeks, mother and father's age, mother and father's educational attainment, and family income—to name a few. Again, all of these are captured at the time of the child's birth.

3.3.2 College data and enrollment process

I use data on the universe of all males and females who were enrolled into a college or university between 1996 to 2001. These data come from the College Entrance Exam Center, the independent government organism attached to the MOE that is responsible for creating, disseminating and grading the Joint-College Entrance Exam (JCEE) and assigning the students to their institution and department. (For simplicity, I will use MOE to refer to both Ministry of Education and College Entrance Exam Center.) Unfortunately, for these cohorts, I do not have data on all JCEE test takers, only admitted students. After cleaning the data, this leaves 381,054 observations. Table 1-Panel B provides the descriptive statistics. Between 1996 and 1997 the cohort size of JCEE-passers more than doubled, as shown by Figure 2. The timing of this increase in JCEE pass-rate is distinct from the expansion of the number of colleges and universities, which it precedes by one year. This differential timing of increased probability of attending colleges versus universities adds another layer of exogenous variation to the policy experiment. I discuss this more fully below.

In 1996, there were 3 percent points more females attending college than males, but this advantage was lost starting in 1997, whereupon the average cohort had the same sex-ratios as the birth cohort. At around 17% urban-dwellers, college-goers during this period reflect the overall urbanicity of the population, as represented by the 15% urbanicity of those in our birth data.

The data contain information on the following: the year for which the JCEE was taken, the institution's name where the MOE assigned them, the name of specific department in the institution where the students were assigned, a specific MOE code number for that year for that institution-department (CD) pair, the minimum weighted JCEE score for the CD, the number of slots the MOE approved for that CD, the number of slots the MOE actually filled when assigning students, and the gender breakdown of those admitted to that CD.

Three important points follow from the above list of variables. First, since all CDs are capacity constrained by the MOE, the minimum score for each CD represents the JCEE score of the last person assigned to that CD by the MOE. Second, despite concerns that the overall quality of the newly expanding intake would go down, Table 1-Panel B shows that there was a general uptrend in the yearly average of minimum scores for the period 1996 to 2001. Since I only have minimum cut score for each CD, I expect that the yearly uptrend in the minimum represents a lower-bound estimate of the larger uptrend in yearly average uptrend in JCEE scores.

Using the JCEE year and the MOE code, I am able to also link these data to MOE records on each institution regarding its history, location, and current name. The historical data I have access to contain the year when the institution was established; the years it was upgraded from a junior college to college or college to university or both; the year it became a national-level institution; and the years it changed its name. Lastly, it contains the MOE's unique institution identification code that I use to track institutions over time and as they change their names. These historical data are also useful for constructing one of the key treatment indicators, discussed below.

There are several steps in the process by which high school students are assigned to their CDs. During the first week of July every year, high school students who hope to attend college sit for the JCEE on the same 3 days in their high school. Students have a range of subject-specific tests they must choose from, depending on the educational track they would like to pursue in college. Some tracks, like Business, only require 3 subject-tests. So since each of these tests is worth 100 points, the scores to get into a business program will be on the lower side. Other more difficult tracks, like Engineering, require 7 subject-specific tests; thus, the JCEE scores to get into these tracks will be much higher. If a student doesn't take one of the required subject-specific tests, then even if the total score is high enough, she will not be granted a slot in that program. Other important details regarding the subject-specific nature of the JCEE remain, but are less relevant to this particular analysis, and thus are ignored.

Once a student has taken her JCEE, her test is sent to the MOE to be graded. The MOE returns to each JCEE-taker her own raw score with the minimum, mean, and maximum raw scores for the JCEE that year. Along with these scores, the college-hopeful is sent an 'aspiration card' to rank, in order of preference, the specific CDs she would prefer to attend, given her preceding choice of which field of study. These are returned to the MOE. The MOE uses this information to allocate students to CDs, starting with the student with the highest weighted score and moving down until all slots are filled. CDs must enroll those students they are allocated, and students must attend where they are assigned or wait until the next year to try again. Importantly, the student's incomplete information, the binding nature of the MOE's decision, and the MOE's central role in the allocation process—all work similarly to (and I believe better than) Dale and Krueger (2002)'s attempts to remove students' selection into institutions and institution's selection of students.

3.3.3 Earnings Data

The administrative data I use on earnings is from Taiwan's Bureau of Labor Insurance for the years 2000 to 2005. The average size of each cohort is around 6.5 million individuals. It is an employer-employee matched data set, which includes the monthly earnings for that year and the employer's unique ID number. The major advantage of the data is the sheer number of observations; the data contain nearly the complete universe of all full-time workers in Taiwan between 2000 and 2005.

Table 1-Panel C provides the average monthly earnings by college-intake cohort for those whose first observation appears in that year. Of the 381,054 college-goers in our education data, I achieve a match on first-year earnings for 353,456—a match rate of 92.8%. About

10% of public sector employees, 3% of farmers, and 5% of self-employed, and all unemployed are not in this data. The former three maintain a different form of public insurance that pre-dated the near universalization of the Labor Insurance program in 1995 (Chou et al., 2003). The latter has no labor insurance because they do not have a job.

The data on the monthly earnings, measured in nominal terms and the local currency of New Taiwan dollars, are recorded in December for every year. The advantage of having monthly earnings from December is that I can capture first-year earnings soon after graduates complete their learning in June of each year. The second advantage that I enjoy over monthly earnings data recorded in the typical survey is that I record all observations on the same month of the year, which protects against the noise induced by business-cycle seasonality in monthly earnings. Using an independent data source on average monthly earnings of this period (see Figure 4), I see that December earnings are representative of other months in the year, except January. January's significantly higher average for monthly earnings likely captures the fact that yearly bonuses are paid out during Lunar New Year, which occurs during January.

A disadvantage is that I have earnings data for only 1 of the 12 months, but this is not unlike most analyses which observe only one observation per year of earnings. We, therefore, treat the extra knowledge I have about when exactly the data were observed as an overall improvement. A second limitation of this data is that monthly earnings data were measured somewhat roughly. The upper limit in the records is 42000 NTD, which is sufficiently high for most first-year earnings of new graduates. And the lower limit is 11,000NTD, which again is well below the average first-year earnings. Lastly, monthly earnings are listed in clumpy intervals, not in single-dollar units. I interpret this as working against my favor by increasing classical measurement error and biasing my estimates towards finding a null effect.

The summary statistics in Table 1-Panel C bear the fact that the average first-year earnings falls well within the upper and lower limits of the measurement of monthly earnings.

The average first-year monthly wage ranges from around 22,000 to 35,000. The wide range over the wage years (2000-2005) can be explained by two phenomena, I believe. First, since males must serve 2 years of national service, I usually do not observe them in the earnings data for the wage year of their graduation, nor the year after. And since females tend to pursue higher education that leads to professions with lower salaries, the first two years of average earnings are characterized by this. Reading from left-to-right by college-intake cohort, two years after graduation we see the average first-year earnings increases by about 3000 NTD, which is what we would expect with the males entering the labor force for the first time. This delayed entry into the labor force for males and females requires that all regression analyses be run separately for males and females. Second, the higher first-year earnings we observe in later wage years (2003-2005) for the earlier educational cohorts (1996-1997) likely represent those who pursued further postgraduate studies. Yang et al. (2010) do show that following the expansion of higher education at the undergraduate level, there was a concomitant increase in postgraduate studies as well.

Finally, I use first-year earnings because it is the most sensitive measure of employers' responses to the expansion of higher education in Taiwan. If employers were using signals like 'university' and 'college' to gauge the average productivity of the graduates from any given institution, then the expansion may have distorted the precision of these signals, and thus the variance in first-year earnings should follow a similar path as the variance in these signals.

3.4 Identification

I utilize several features of Taiwan's schooling system and access to uncommonly rich and large datasets—birth, educational, military, and employment—for the universe of individuals born from 1978 to 1983. The primary objective of this paper is to estimate the signaling effects of higher education, net the human capital effects. In order to do this, I first establish that this policy change increased the probability of enrollment into a college or university. Accordingly, I use both regression discontinuity (RD) and nonparametric regression techniques. When identifying the main signaling effects, I construct fixed-effect (FE) specifications in difference-in-differences (DID) and event-study designs and rely on relevant policy details to provide a causal interpretation.

3.4.1 Effects of Month-Year of Birth on Probability of Attending an Institution

I estimate the impact of the increased access to any institution on the birth cohorts 1978 to 1983. Since these estimates are not the main focus of this analysis, I provide the results in this section. The first strategy I deploy is a school-entry age RD design, featuring year-month of birth (as in born July 1980) as the running variable and various probabilities related to attending an institution as the outcomes of interest. Such RD entry-age designs have been used to great effect recently in the education economics literature (Oreopoulos, 2006; McCrary and Royer, 2011; Dobkin and Ferreira, 2010).

Taiwan's rules for starting first grade require that a child be 6-years of age by Sept 1st. If any child is not, then she must wait until next year. Applying the rule to this college context, a child born before Sept 1st, 1978 should have taken the JCEE in July 1996, before the expansion. A child born on or after Sept 1st, 1978 but before Sept 1st, 1979 should have taken the JCEE in July 1997, the year the JCEE pass-rates were increased such that there was increased access to higher education generally. Finally, one born on or after Sept 1st, 1979 would have taken the JCEE in July 1998, the first year that colleges were allowed to upgrade to universities.

Thus, right at September 1978 I should see a jump in the probability of attending any college or university. At Sept 1979 we should see a jump in the probability of attending a university instead of a college. The RD specification that I use is shown in model (3), where $(Sept1979)_i$ is an indicator for whether a child is born on or after Sept 1979, which implicitly includes Sept 1st. The parameter $f(ym)_i$ is a flexible function of each individual's year-month of birth. Hahn et al. (2001) show that with a flexible enough specification, $f(ym)_i$ is non-parametrically identified around the cutoff. I use both quadratic and cubic functional forms; both provide similar results. I present results with cubic. Table 2 provides the estimates.[5]

$$Pr(Y_i) = a + Sept1979_i + f(ym)_i + e_i \tag{3.1}$$

I consider three outcomes of interest for $Pr(Y_i)$. The first $Pr(HEI)$ is the probability of attending any form of higher education. This is a lower-bound estimate of the more policy-relevant estimate $Pr(HEI|JCEE)$, the probability of attending an institution, conditional on having taken the JCEE. Unfortunately, considering the limitations of that data, the first-best approximation I expect to offer is $Pr(HEI|HS)$, the probability of attending an institution, given attendance at high school. This still would be an improvement.

For now, I acknowledge the measurement error and simply interpret the lower-bound estimate as the chance anyone in the birth cohort had at attending an institution, given her year-month of birth. Making these acknowledgments, I find that those born on or after Sept 1979 enjoyed a 4.8 percentage-point increase (ppi) in the probability of attending an institution by having had to wait another year to start school. The second outcome $Pr(aft1997|HEI)$ focuses on the probability of having taken the JCEE in 1998 or after, given that one ended up at an institution, meaning I can observe the individual in the data. This outcome is an intermediate-stage outcome. My causal story has been that children born

[5]Because our data start at Jan 1978, we do not focus on the Sept 1978 discontinuity in the RD specifications.

on or after Sept 1st 1979 should have taken the JCEE in 1998, if they were going to take it at all. For this I restrict the sample to only those who attended any institution; Columns 2 and 5 in Table 2 show that the likelihood of taking the JCEE in 1998 instead of 1997 was around 50% if one was born on or after Sept 1979. Since I have JCEE data only for those who ended up attending some form of institution, I am unsure whether this is an over- or under-estimate in the true probability of having taken the exam at the birth month Sept 1979. This could be corrected for males using data merged with military conscription data that collects years of schooling and then imputed for females. But generally, I believe the trend breaks around Sept 1979, shown in Figure 6, provide compelling reasons to assume that even with measurement error the increase in probability of taking the exam in 1998 was substantial.

The third outcome $Pr(Univ|HEI)$ measures the probability of attending a university instead of a college, given that the student went to an institution. This is the main outcome of interest. For this I again restrict the sample to only those who attended any institution; Columns 3 and 6 show that the ppi in probability of attending a university over a college was 9.5 for males and 8.2 for females who were born on or after Sept. 1979. If we were to construct a Wald estimate from this — that is, $\frac{Pr(Univ|HEI)}{Pr(aft97|HEI)}$ — the probability of attending a university over a college, inflated by the probability of having taken the JCEE in 1998 or after, we get a 20.45 ppi for males and 15.30 ppi for females. All the above estimates are statistically significant at 1% level.

Though the RD estimate is not the main focus of this paper, I follow standard practice and provide the common robustness check, which will be useful later. I test for the smoothness of other characteristics across the Sept-1979 point of discontinuity in the distribution (McCrary, 2008). For this, I refer back to the summary statistics displayed in Figure 3. Common characteristics that are explored in other research are cohort-density, gender ratio, family income, and maternal age. Figure 5 includes other less common, but equally interesting characteristics that the data contains, such as whether or not the individual was the eldest, whether or not a parent worked for the government, urban status, and whether or not the mother held a high-school diploma. All the characteristics in both Figures 4 and 6 are smoothly distributed across the Sept-1979 thresholds where we see discontinuous changes in the probability of attending higher education. These smooth trends in other characteristics bolster confidence not only in the RD estimates I presented above, but also the non-parametric estimates I describe below.

The RD estimates were an overly strict focus on the Sept-1979 discontinuity. Given the magnitudes of this policy change, I should expect that the general equilibrium effects are of greater import than the partial equilibrium ones. Therefore, the second methodology I deploy uses nonparametric regression again of $Pr(HEI)$, $Pr(aft97|HEI)$, and $Pr(Univ|HEI)$ to see the whole distribution. Figure 6 visualizes the coefficients for these regressions. The $Pr(aft97|HEI)$ and $Pr(Univ|HEI)$ impact are visually striking. Right around Sept 1979 there is a large, sustained jump of about 22 ppi for $Pr(aft97|HEI)$ and 11 ppi for $Pr(Univ|HEI)$ right at Sept 1979. For $Pr(Univ|HEI)$, this level-rise is already on top of the new base that was established after the Sept 1978 jump of about 5 ppi.

Over the whole distribution of year-months of birth we see a very close match between the pass-rates in Figure 2 and $Pr(Univ|HEI)$ in Figure 6. The major trend breaks that occur around the Sept-months in 1978, 1979, and 1982 represent the separate waves of the policy being implemented. The Sept-1978 jump results from the overall increase in passing rates for the JCEE that was instituted in 1997. The jump at Sept 1979 results from the differential increase in the number of universities relative to number of colleges, which occurred in 1998 because of the upgrading already discussed. The estimates of $Pr(Univ|HEI)$ are remarkably flat after Sept 1979, which reflects the flat trend in passing rates we see in Figure 2. I attribute the flatness to the number of JCEE-takers increasing at about the same pace as the number of seats at universities for the 1997 to 2000 JCEE cohorts. Lastly, the jump around Sept

1982 reflects the second wave of upgrading institutions that occurred in 2001 (see Figure 1) and shows up in the passing rates in Figure 2.

The cyclical trend that we see in $Pr(HEI)$ over the period from 1978 to 1982 highlight the fact that those children born right after Sept 1st each year enjoyed some sort of age-advantage which increased their probability of attending an institution. This age-advantage is a complaint Currie (2009) brings against such birthdate RD designs, specifically for the McCrary and Royer (2011) case. Interestingly, in my case, this cyclicality is not replicated in the $Pr(Univ|HEI)$ trends. I argue that this suggests the underlying mechanism generating the cyclicality in $Pr(HEI)$ is the age-advantage junior high school students enjoy when taking the centralized national high school entrance exams every year. Since high school is the gateway to taking the JCEE and since $Pr(Univ|HEI)$ is conditioned on having taken the JCEE, the cyclicality is netted out. Therefore, Currie's concerns are not applicable in this case.

This discussion of the nonparametric estimates underscores the importance of including the general equilibrium effects in the analysis. Both the RD and nonparametric estimates agree in magnitude and timing of the effect, and provide compelling evidence that this higher education reform significantly, substantively increased the probability of attending an institution, and specifically a university over a college. I now move on to the main focus of my analysis: testing for the signaling effects of university over college, net the human capital effects.

3.4.2 The Signaling Value of School Reputation

Similar to Tyler et al. (2000)'s work with the GED, this paper tests for the signaling value of education, independent of human capital effects. We explore this within the context of previous theoretical (MacLeod and Urquiola, 2011) and applied (Dale and Krueger, 2002, 2011; Tao, 2007) work on the returns to school reputation.

Unlike most papers, I couple a novel policy experiment with the universe of individuals affected by that policy and for whom I have a rich set of variables to measure impacts and control for confounding effects. Tyler et al. (2000) highlights the central importance of uniting the right data with the right policy variation: they write, "Ideal data for identifying the returns to a signal would contain exogenous variation in signaling status among individuals with similar levels of human capital" (p.432). I believe I have met these requirements. Below I justify this belief. First, I justify the comparison to Tyler et al. (2000). Tyler et al. use the fact that the GED is a national test but the scores needed to pass the GED are different at the state level. Hence, for two students each in a different state but scoring the same grade on the GED, one could have passed the GED and the other failed. Using the same GED score to control for human capital differences, but the variation in who got the GED by state, they causally estimate the signaling value of a GED. Once they enter the labor market, the GED holder will receive a higher wage, despite not having any higher human capital, as measured by the GED.

I focus instead on the differential returns to the signals of 'college' versus 'university.' As Figure 1 (above) illustrates, the higher education expansion was accomplished by providing a centralized method for colleges to upgrade into universities.[6] Importantly, institutions were allowed to upgrade as long as they could document that they had met the minimum standards of the next higher status institution for some years before applying to be upgraded. The most basic interpretation of those requirements means that for at least two years (one before application and one during the application process) before the college expansion (e.g., 1996-1997 for those HEIs that upgraded in 1998), there existed some colleges that were arguably as good as the lowest reputation universities. Despite this essential equality, for

[6]MOE Order No. 85088802, `http://english.moe.gov.tw/content.asp?CuItem=8205&mp=1`

bureaucratic reasons these institutions and their graduates were not allowed to use the more prestigious label 'university.' These graduates from higher-reputation colleges and lower-reputation universities are each other's counterfactual, before the expansion. After the expansion began, the alum from colleges prior to their upgrading became the counterfactuals for new graduates after the colleges were ranked as universities.

In 1998, the first 13 colleges upgraded to university status changed their name to include 'university' and were assigned university students during the JCEE student allocation process. Using this group of upgraded colleges that upgrade and cohorts of students enrolled just before and after the switch, I can effectively control for relevant differences between colleges and universities over this period, down to the very department the students were in. Therefore, I recover a causal estimate of the returns to reputation, net human capital effects. This is a fundamental element of my identification strategy to test for signaling effects.

With the fine-grained and extensive data I have, I can run a range of FE specifications that control for the year of birth (B_i), the specific year the student took the JCEE (J_i), the specific years for which I observe monthly earnings data (W_i),[7] and even the extremely strict specification of including an indicator for the specific institution or CD (CD_i) the individual was enrolled into, as shown in model (2). The vector X_i includes controls for type of institution (private, public, normal, or top-6), parental characteristics at time of an individual's birth (education, log-income, county-of-residence), and an indicator for whether the major was a special course requiring more than 4 years. Standard errors are clustered at the CD-year level to capture the fact that the MOE allows each institution to make enrollment decisions at the department level. Specifically, the MOE allows each CD to negotiate its own intake size and minimum cut-score for that year's intake.

I run separate regressions for males and females. I do this because males must perform

[7] Including both J_i and W_i implicitly controls for years of work experience, which would be relevant in any earnings growth estimates and is common in the Mincer model of returns to schooling.

mandatory military service for a standard 2-year period sometime after their 19th birthday. Those who attend post-secondary education may defer their service until they complete their schooling. Because the timing of their conscription period could vary, this could affect the timing of their entry into the labor market. This is systematically different from females, who mostly enter the market immediately following graduation.

$$Y_i = a + J_i + B_i + W_i + CD_i + treat_i + treat_i * center_i + X_i + e_i \tag{3.2}$$

I use two DID variables for $treat_i$, constructed from the data: $UNIEVER$ and $UPGRADE$. $UNIEVER$ identifies whether or not the individual enrolled at an institution with the label 'university' in the name at the time she took the JCEE.[8] This information is taken from the College Entrance Exam Center data. As a second measure, $UPGRADE$ uses a difference source of data the MOE records for the year an institution gained university status. Both of these treatments work like DID estimators, because they switch from 0 to 1 in the years following the college attained university status between 1997 and 2000. If the institution had been a university before 1997, $treat_i$ remains 1; conversely, if it achieved university status after 2000, then $treat_i$ remains 0. The J_i fixed effects supply the trend-control that is usually more roughly approximated by a before-after dummy in most DID designs. The interaction with $center_i$, which measures the years between when an institution upgraded to the first year I observe monthly earnings, captures the changing effect the new signal has on the labor market—a parametric approximation to what I observe more clearly in the event study analysis. Lastly, Y_i measures log of first-year earnings. Therefore, the point estimate on $treat_i$ is the ppi in monthly earnings one can expect from getting into university instead

[8] All institution names taken from the JCEE were current to the exam year for which we have the data. The names in the JCEE data are in Chinese; I use Google Translator to translate all the names. All the names were translated over a short period of time, and the names for universities and college are common and translate in the same way. They are "university" (*daxue*), "colleges" (*xueyuan*) and "junior colleges" (*zhuanke*).

of college. Both treatment variables provide similar estimate. Table 3 displays and section 5.1 reviews the DID results.

I have provided various forms of evidence that support my causal story. There has been some concern that a DID estimate can suffer from serial correlation, which can lead to an over-rejection of the null-hypothesis or a higher probability of finding an effect (Bertrand et al., 2004). Bertrand et al. (2004) show how DIDs with limited variation over long panels can tend to suffer from serial-correlation, causing the estimated standard errors to be overly tight and t-statistics to be biased upward. To control for this, I narrow the sample down to using only those individuals who attended a college before and after it upgraded between 1997 and 2000. By using only the institutions that upgraded, I work with only those institutions that do experience variation in treatment, thereby removing concerns regarding whether or not the standard errors correctly estimated . This is an extreme test because it reduces both my sample and the number of CDs that show up and because I cannot now rely on other older universities to bolster the average strength of the signal. The estimates from these regressions can be interpreted with greater confidence in their precision.

In the fourth and final research strategy, I perform an event study analysis (model 5) to complement and extend the standard DID framework. For this strategy, I restrict the sample of only those individuals who attended a college before and after it upgraded to university status between 1997 and 2000. In addition, because an event study is a nonparametric technique, it is a more flexible version of my DID. The flexibility makes the analysis stricter on the data, because it does not force linearity on the coefficients of interests, a linearity that can sometimes increase the likelihood of identifying significant results in DID. Moreover, beyond being just a corrective, an event study affords us a perspective on the changing trends around the policy change, not just the point estimate at the time of the policy change. Taken together, these two advantages make the event study a robustness check on the already strict DID specification I presented above as well as a move toward considering the general equilibrium outcomes that this policy caused.

$$Y_i = a + J_i + B_i + W_i + CD_i + bef3_i + bef2_i + aft0_i + aft1_i + aft2_i + aft3_i + X_i + e_i \quad (3.3)$$

All results are the nonparametric coefficients from $bef3_i + \ldots + aft3_i$. The base year is the $bef1_i$ (-1 in Figure 7), and is left out of the regressions. The dummy variables are centered in relation to the year the college upgraded to university-status (aft0i, or 0 in Figure 7). All aft-coefficients are significantly different from the null at the 1% level, the *before*-coefficients at the 5%. These log-earning trends capture the increasing reputational benefits of attending an upgraded college now that it is a university.

3.5 Results and Discussion

All the results we show are causally identified for the 6 birth cohorts affected from 1978 to 1983, providing a general equilibrium estimate of the monetary returns to attending a university over a college. Using only the colleges that upgraded, we estimate the reputation effects, net of any human capital effects. I believe that I have achieved the data/quasi-experiment combination that Tyler et al. (2000) described as the ideal test for the effects of signal.

Moreover, beyond my contributions to the signaling literature, the analysis is generally interesting within the causal inference literature because I estimate a LATE parameter which differs from the normal LATE estimates in two ways. First, this LATE parameter is situated at the upper tail of the educational distribution, instead of the more common lower tail. Most causally identified estimates of the returns to schooling estimate the effect of the marginal student who is considering dropping out early but must go on to finish one more year of schooling. This extra year of schooling often still leaves the student far from completing a high school diploma. Second, because of this upper-tail quality, the 'compliers' in my study (those who now can get into an institution or, in particular, a university) are, in fact, very much like the 'always-takers,' to use Potential Outcomes language (Rubin, 2005). In most Asian countries at least, college is the dream of almost every high school student; yet the limited supply prevents it. Therefore, in a policy context that seeks to balance the twin concerns of maintaining a meritocratic yet equitable society, this is the subpopulation one wants to target.

3.5.1 DID results

Table 3 provides the main DID results for the different FE specifications used in this analysis. Because I have two separate treatment variables ($UNIEVER$ and $UPGRADE$) and separate regressions for males (m) and females (f), I have chosen to show estimates with the general set of fixed effects (J, W, and B) and the full set of controls included. As a general rule, including the controls in the regressions increases the point estimates by less than half a percentage point and leaves the statistical significance unaffected; hence, the tables below are a fair representation of the overall analysis. The only variation in specification I show here involves the type of fixed effects that I am interested in, that is, those related to the specific CDs my sample are enrolled at.

A quick preview of the estimates shows that the returns to log-wages for males range from a 6.2 ppi to 3.5 ppi across the different specifications. These are quite substantial returns, considering that both university and colleges provide four years of post-secondary education and terminate in a Bachelor's degree. This is less, however, than the 10 to 19 percent estimates that Tyler et al. (2000) show for passing a GED versus not, controlling for scores on the GED. For females the range is smaller and the average lower, encompassing 2.9 to 4.3. It is a fair generalization to make of my results that across the specifications the point estimates for females average about 30 percent lower than those of males.

Focusing on the different fixed-effect specifications, when the specific institution is controlled for, the point estimates are larger, as may be expected. This level of specificity in the fixed effect model goes beyond most analyses. The point estimates on $UNIEVER$ are noticeably larger relative to those on $UPGRADE$. Conceptually, we may prefer the estimates on $UNIEVER$ if we believe that the JCEE data from the College Entrance Exam Center's records are most accurate. We would prefer the estimates on $UPGRADE$ if we believe that the MOE's records are more accurate. The JCEE data for each year contain information on over 1200 CDs per year, so the data are much finer, but there are more data points and thus greater potential for error. The MOE's data have information on over 130 institutions, which is much rougher but also more likely to be error-free. Since the above description does not lend itself to preferring one over the other, I present both for robustness. Generally, they both agree in magnitude and significance across the different specifications.

By switching to the CD fixed effect, I remove the time-invariant characteristics of specific department that each student is in. This implicitly controls for the idiosyncratic details that make, say, one Math department different from another at a different institution; but it also controls for what makes the Math department at National Chengchi University different from the Political Science department there. Thus, by using the CD fixed effects, I am controlling for both the institutional and the curricular aspects that make one department different from another. The organizational change literature on higher education describes institutions as possessing much institutional inertia. Thus, I believe that using the quite stringent CD fixed effects captures a great deal, even though it absorbs only the variation due to time-invariant characteristics.

Making these much more stringent claims on the data do reduce the point estimates on both $UNIEVER$ and $UPGRADE$ from what they were when only the institution fixed effect was used. For both $UNIEVER$ and $UPGRADE$, the reduction is about 1.5 ppi of the effect—a non-negligible 25%. But quite interestingly, this is only the case for males. When looking at the females, their estimates are actually about a half a percentage point higher. Focusing for now on just comparing male-female estimates within the CD specification and within the respective $UNIEVER$ and $UPGRADE$ treatments, I see that they are fairly well balanced now. Whatever was causing the nearly 30% lower ppi for females before under the institution fixed effects was dealt with by controlling also for the specific department within the institution.

Moving on to my strictest specification thus far, I use only those individuals who attended a college before and after it upgraded to university status between 1997 and 2000. As remarked above, this drops both individuals and CDs, thereby reducing along two dimensions my statistical power to detect an effect. Nevertheless, if we believe that these institutions are those most affected by the policy change, then focusing only on them is useful and necessary.

When I use only those colleges which upgraded to university-status between 1997 and 2000, the results are qualitatively similar in direction and statistical significance to the results from the larger sample who attended any institution. Unexpectedly, the gender differential in wage premia for university graduates relative to college graduates has returned. In the case of $UPGRADE$ as the treatment variable, the effect is even more pronounced now, with males enjoying a wage premia twice that of females. Though I do not have the space to pursue it further here, I show here that males appear to have benefited the most from the expansion.

3.5.2 Event Study

In the above, I show the cohort-average effects of the expansion, focusing all the way down to only those students in institutions that upgraded. I now switch perspective to test the timing effects of the outcomes, focusing again down to only those individuals in institutions that upgraded. In the event study design, I have re-centered each JCEE cohort relative to the year the institution they attended upgraded to university status, where zero on the x-axis is the first cohort to enter the newly upgraded institution. The counterfactual in these regressions plays each cohort off the other—both within/between upgrading institutions and before/after institutions upgraded.

Table 4 has the estimates for males and females, and Figure 7 displays their separate trends. The coefficients displayed in Figure 7 are taken from the bef- and aft- dummy variables in model 3.[9] Four patterns are evident from both Table 4 and Figure 7. First, looking at Figure 7, I interpret both the direction of the trends and the timing for their crossing from negative to positive effects to validate my signaling story. Second, we see that the average effects estimated off $UNIEVER$ and $UPGRADE$ in the DID specifications (model 2) in section 5.1 are fairly representative of what is seen at period 0, the first cohort of students who graduated with a university-BA degree instead of a college-BA. However, the increasing trends extending into periods 1 and 2 — or the second and third years — suggest for both males and females that employers were increasingly rewarding these upgraded university-BAs, possibly as their faith in the reliability of the signal increased. A similar interpretation could be given to the interactions of center with $UNIEVER$ and $UPGRADE$, which is another example of consistency between my methods.

Third, looking now at Table 4, all but one of the aft-effects are statistically significantly different from the null. Even the $before$-coefficients are significant at the 10% level, at least. Looking two periods away in either direction, I can say with 95% confidence that the estimates are statistically different from each other, implying that the upward trend is a valid interpretation. These estimates are substantively large and suggestive that the there was a preference for these university graduates over those from their alma mater in previous

[9] All coefficents are relative to the base year, which is the year before the colleges were assigned university students by the JCEE; thus, $bef1$ is not included in the regressions or in Table 4.

cohorts and their contemporary college graduates. Fourth, I see again that males enjoyed higher returns for longer, though they eventually return to the same level as females. Still, their confidence intervals do overlap, though neither gender's point estimate itself lies within the other's confidence interval.

Taking all this evidence together, I believe my causal story is validated. The estimates suggest that the signaling value of university's reputation over a college's, even within the same institution, is substantial, ranging from 2 to 6 ppi for the first cohort and higher for later ones. Indeed, acknowledging the various methods and the increasing levels of rigor in research design, my story has held up well. And yet, I put the story to one last test.

3.5.3 Controlling for time-varying department level characteristics

So far I have relied on policy details of a novel natural experiment and very detailed individual-level data to control for time-invariant characteristics using CD fixed effects. I have argued that in order to upgrade, the college had to document that they have met the minimum requirements of a university for at least two years prior to applying. In effect, they were acting like universities before they were allowed to be called one. This alone is sufficient to estimate a signaling value of these colleges upgrading to university, assuming the experiment was as in internally valid as I have shown and believe it to be. Moreover, my intensive fixed-effects and restricted-sample regimes have done well to net out the time-invariant characteristics that could have explained some of the variation in log-wages. Both of these are a contribution to the literature. I extend them one last time to handle the latest evidence in the 'better schools' literature, regarding behavioral responses to school quality.

Pop-Eleches and Urquiola (2012)'s survey of principals, parents and students shows that people respond to changing school quality. From this, they argue that more care should be applied to accounting for these changes in empirical estimates of school quality. This is quite hard to do without the benefit of right data. Thankfully for us, the Taiwan MOE makes available a very detailed accounting of each CD for each year. Specifically, for this context, because each CD is allowed to negotiate with the MOE directly regarding its intake size and minimum cut-score, I show results for these.[10]

Figure 8 shows the effects of an institution upgrading on CD-level minimum scores and intake sizes. Both scores and intake sizes make up two of the most commonly researched school quality measures in the human capital literature. Both are summarized in section 5.2. An increase in the minimum cut-score intuitively (but not necessarily) means that there was a concomitant increase in the average peer quality, the main mechanism in the recent 'better schools' or 'peer effects' literature. Further, a reduction in intake sizes implicitly means smaller class-sizes, possibly the most heatedly debated school resource. To be explicit, since capital investments in school infrastructure is quite inelastic in the short-run (Urquiola and Verhoogen, 2009), a reduction in the cohort size means that class-sizes will be smaller, following the general pattern of evenly distributing students across classrooms.[11] Thus, I believe that my ability to address these two issues makes a significant contribution in estimating the signaling value of education, net any human capital effects.

I replicate my previous event study design, but now use minimum cut-score and intake size as the dependent variables. Looking at Table 5, Columns 1-4 I can see that there, indeed, was an increase in the minimum cut-score and a decrease in intake-size for departments as their institution upgrades. Importantly, the direction of these trends breaks along when the timing of the policy changed occurred. These pieces of evidence matches what the causal

---

[10] With Taiwan's low-tuition policies, institutions were not allowed to set their own tuition levels without first getting them pre-approved and justified (Tsai and Shavit, 2007). This was not revised until after 2001, after our research period.

[11] Smaller intakes may also mean better teacher-to-student ratios. Because of the tedious nature of the MOE data on teaching faculty, we have yet to merge them with our data. Once we are able to, we believe that we will have covered the three most common quality inputs in the literature, and the only ones that are shown to have any positive, statistically significant effects.

story would predict theoretically from upgrading institutions.

I interpret this as a supply-side response from CDs, meaning that CDs are becoming more selective. An alternative demand-side explanation could be that these effects on the minimum cut-score are being influenced by college aspirants' responses to the new signal of university-status. If higher-scoring college aspirants absolutely prefer attending a university to a college, then because the MOE's algorithm for assigning college aspirants always gives the ones with the highest weighted scores their highest preferred option, it could be that the department's minimum score is never invoked in the student allocation process, even if it were raised after the institution it belongs to upgraded.

This does not explain, however, the downward trend in department-level intake size for these upgrading institutions. Since both the minimum score and the intake size are set temporally prior to students filling in their 'aspiration cards', both could be fairly attributed to behavioral responses by the departments. Since my data do not disentangle these and since it matters little to my need whether it was mainly a supply- or demand-side response, I interpret these effects simply as improvements in school quality.

Smarter classmates and smaller classes are two of the main characteristics any parent or student would use to choose a school. They have consistently been argued in the literature to be mechanisms that improve labor market outcomes. Thus, if including these as explanatory variables in my regressions now causes my estimates of signaling to disappear, what I have thus far been explaining as the signaling value of education would actually just be the human capital value of education.

Hence, I include these in my analysis and re-run both the DID and event study regression to see how much of the variation in log-wages can be explained away by higher peer quality and smaller intakes. Figure 9 and Tables 5 & 6 shows the new estimates for the signaling value of a university degree, net human capital effects. The first cohort (period 0) estimates in Table 5, Columns 5 & 6 are almost exactly the same as what I saw in Figure 7 and Table

4. The steepness in the positive trend is diminished somewhat for those who make up second and third intake cohorts. Yet these are well within the confidence intervals of the estimates in Table 4. Table 6, Panels A and B provides the DID estimates, which are remarkably consistent across the specifications and with the results from Table 3. In this context, the signaling effect of a university degree, over a college degree and net the human capital effects, ranges between 5.0 to 3.9 ppi for males in our sample and 4.9 to 1.3 ppi for females.

3.6 Conclusion

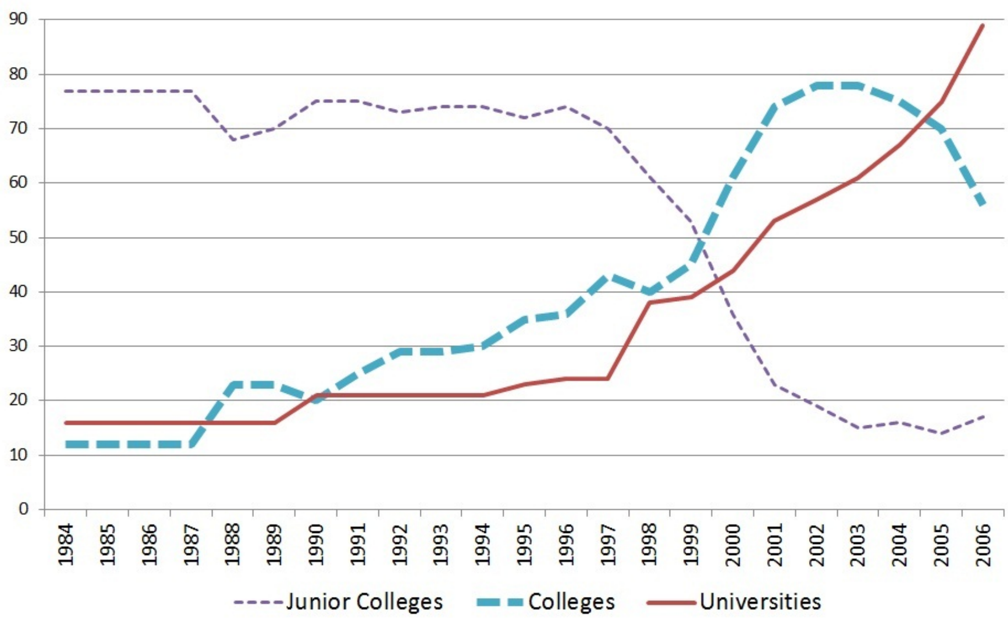
I have used RD, nonparametric, DID, and event study strategies, which impose differing model restrictions on causal story and the sample. Each has complemented the other, and all have developed into a layered, consistent story. Graduates from institutions that carry the name 'university' enjoy a 2 to 5 ppi advantage in first-year log-wages over graduates from institutions that are labeled as colleges. I attribute this advantage to school reputation, which I argue is signaled using two common labels: 'university' and 'college.'

I compare colleges that upgraded to university status within a very narrow time-frame and under strict guidelines which were consistently enforced. The specifics of the institution are that colleges that wanted to upgrade had to prove to the MOE that they had met the minimum standards of a university for some years prior to applying to upgrade. I show from this that even within the narrowly defined set of schools –and their graduates—that were directly affected, the general rule holds. The signal value of a university degree has an average return on log of monthly earnings of about 3 to 4 percentage points. Even after controlling for time-unvarying and time-varying characteristics within the same set of schools, the graduates who were part of the post-upgrade intakes enjoyed a modest advantage in monthly earnings over those graduates who were part of the pre-upgrade intake cohorts.

I acknowledge the potential for behavioral responses in the context-relevant equivalents to average peer quality and class-size, which are the two of the most common research school quality inputs in the literature. Following the advice of the literature, I control for these down to the department level within each higher education institution and for each year. Even after these adjustments, my estimates are robust and consistent with the previous ones in the less constrained models.

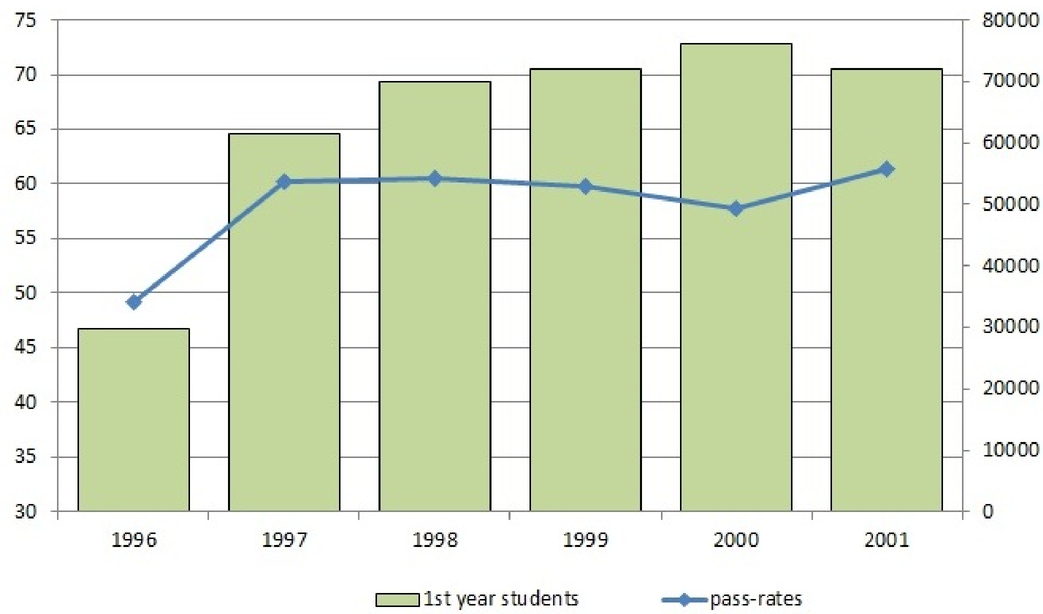
Thus to answer the question: what is in a name? It would appear in this particular context, it is around a thousand extra NT dollars a month. In this paper, I provide strong evidence that there exists a substantial, positive return to the signal of a university degree, independent of any human capital effects that may legitimately be attributed to schools of varying quality. I believe this is one of the first studies to exploit such a union of uncommonly rich data with uncommonly useful exogenous variation in treatment status.

Figure 3.1: Expansion of Higher Education in Taiwan.



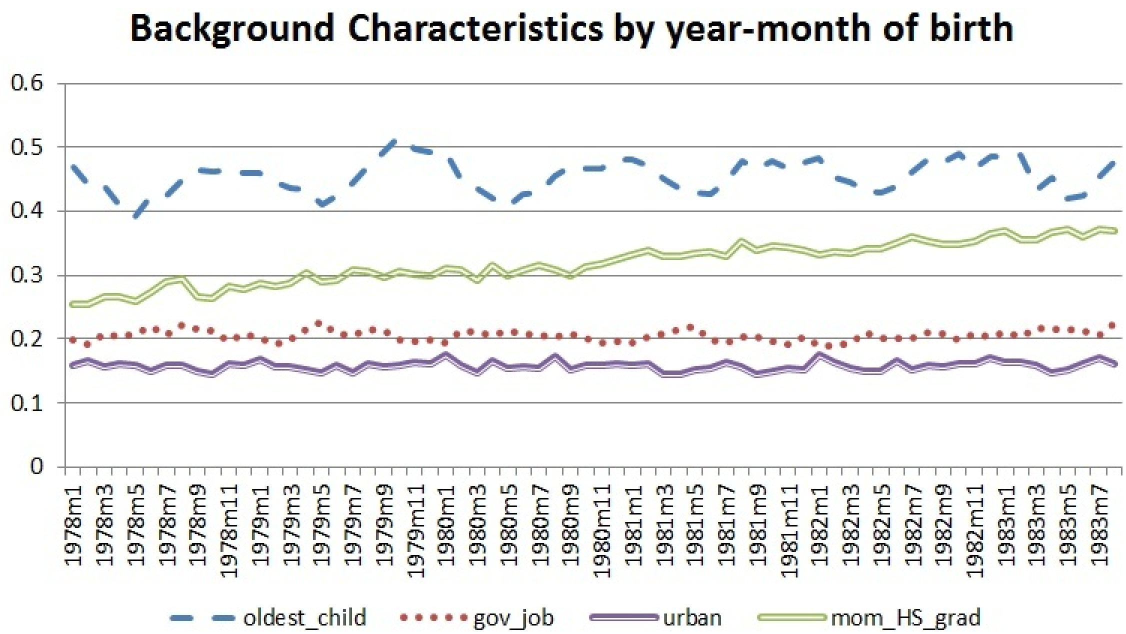
Notes: Unit of analysis in this figure is the institution, not the individual.

Figure 3.2: Number of JCEE takers and passing rates.



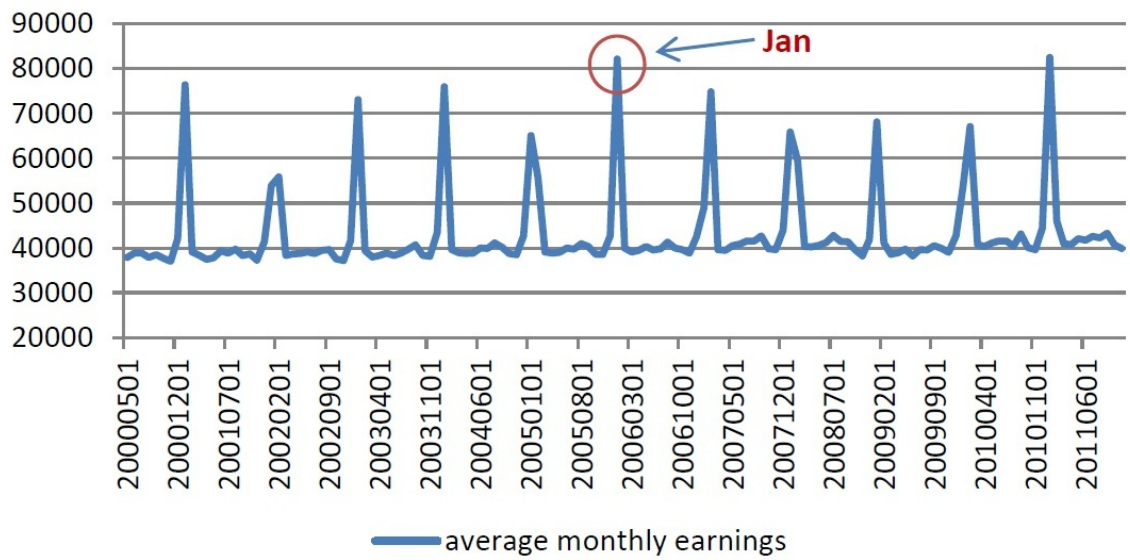
Notes: Source: MOE records. Unit of analysis in this figure is the institution, not the individual.

Figure 3.3: Summary Characteristics by Birth Year-Month.



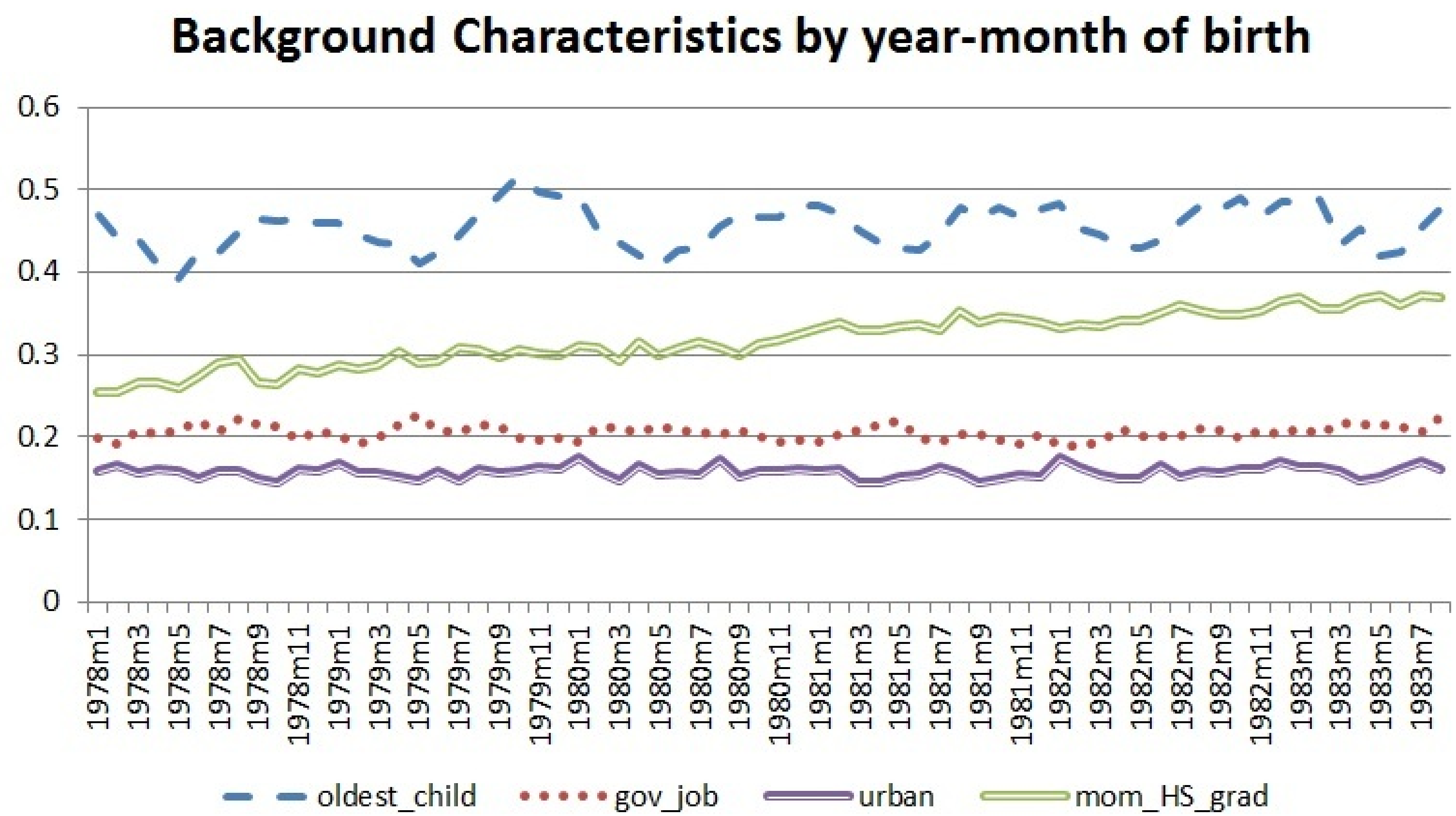
Notes: Author's calculations from the data. Family income represents monthly wages in nominal NT dollars.

Figure 3.4: Average Monthly Earnings in Taiwan: 2000-2010



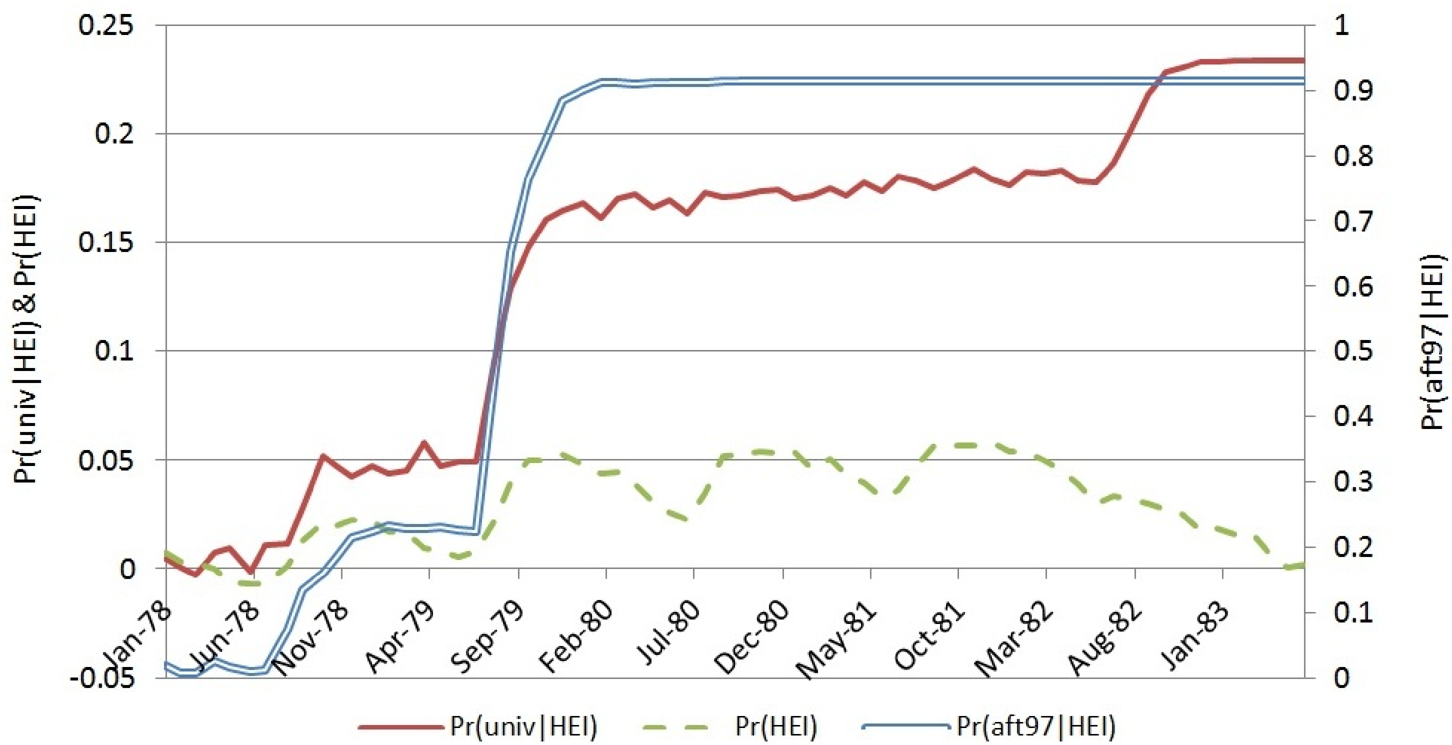
Notes: CEIC data. Monthly earnings in New Taiwan dollars.

Figure 3.5: Trends in Background Characteristics by year-month of birth.



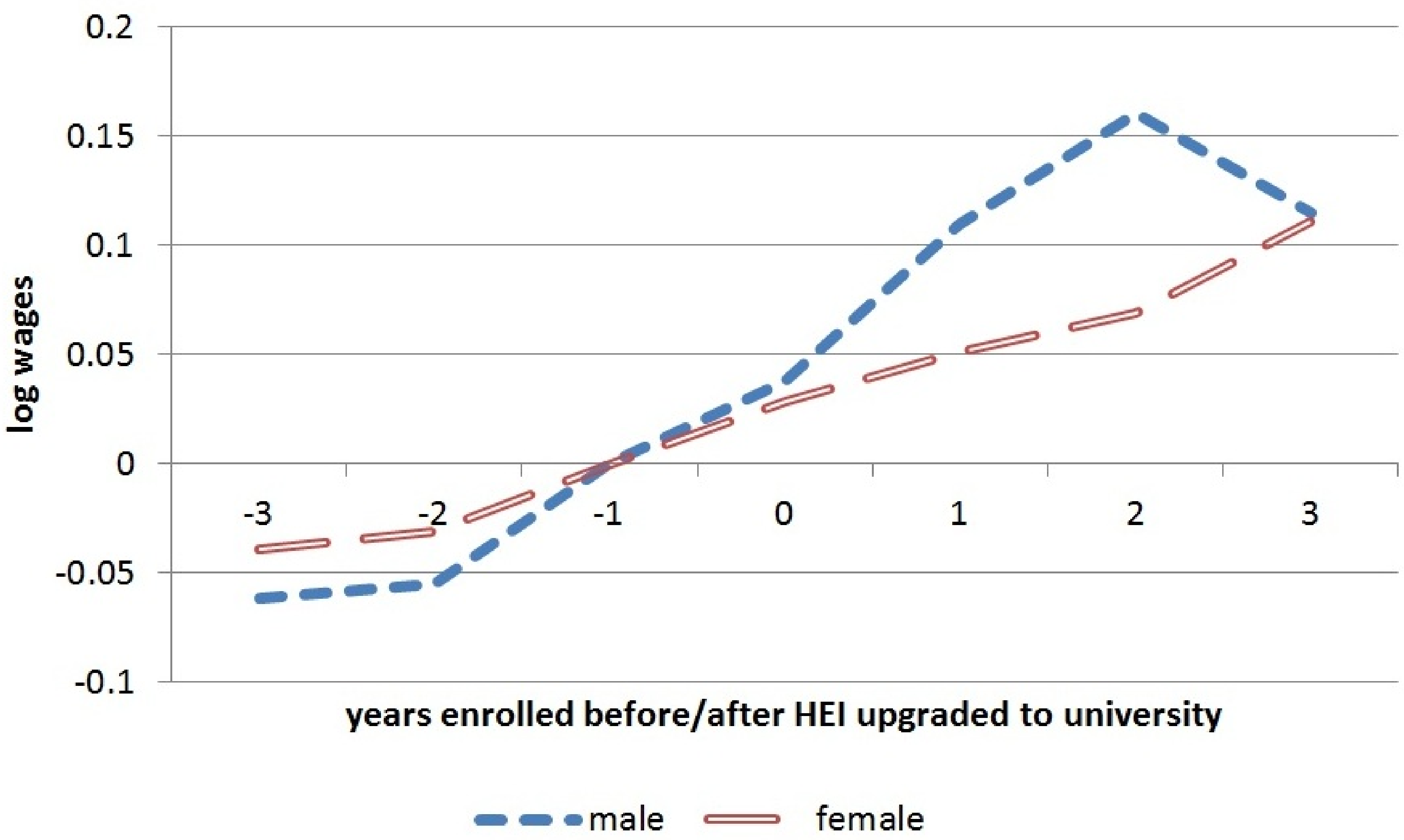
Notes: "Oldest_child" is the percentage who was the oldest child. This inherently presents also only children. "Gov_job" is an indicator for whether one of your parents worked for the government. "Urban" is the percentage who were from the city. "Mom_HS_grad" is the percentage of moms who have a high school diploma.

Figure 3.6: Nonparametric estimates of probability of attending a HEI if born before-after Sept 1, 1979.



Notes: Trends generated from the coefficients off year-month of birth dummies in a nonparametric regressions. Students born after Sept 1979 should have taken the JCEE in 1998, and thus had a differential advantage in getting into a HEI and specifically a university. Pr(HEI) is the probability of attending an HEI; Pr(aft97|HEI) is the probability of taking the JCEE after 1997, given that one is in the sample of college-goers. Pr(Univ|HEI) is the probability that you were enrolled at a university instead of college, given that one is attending an HEI.

Figure 3.7: Signaling value of graduating from a university measured in log wages.



Notes: Year 0 on the x-axis represents (aft0) the first cohort that the newly upgraded university took in. Year 1 (aft1) is the 2nd cohort. The base year is Year -1 (bef1), and so on.

Figure 3.8: Behavioral responses of treated CD's to their newly upgraded university status

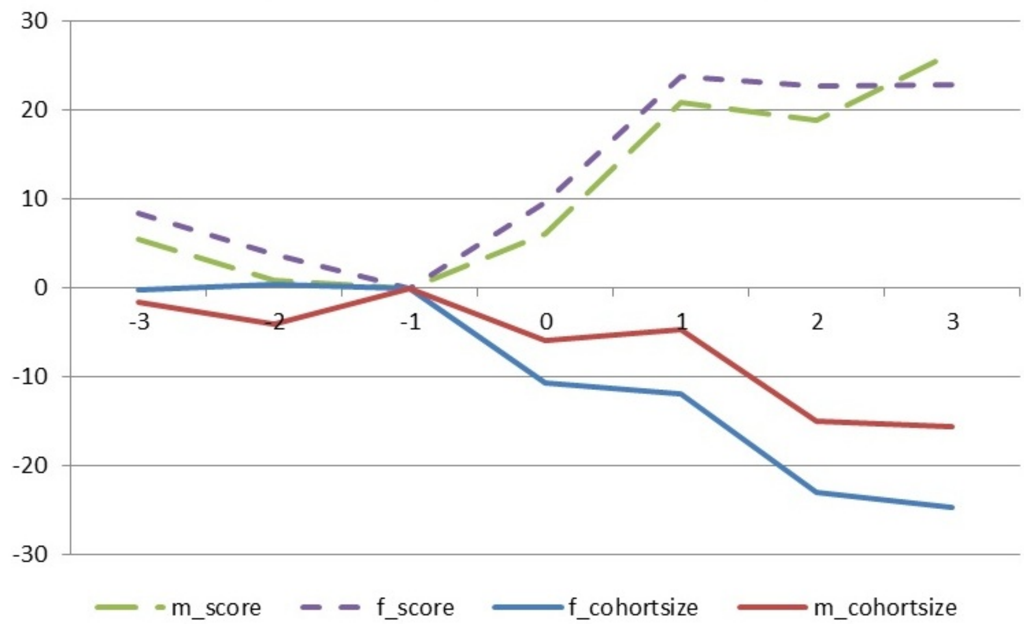
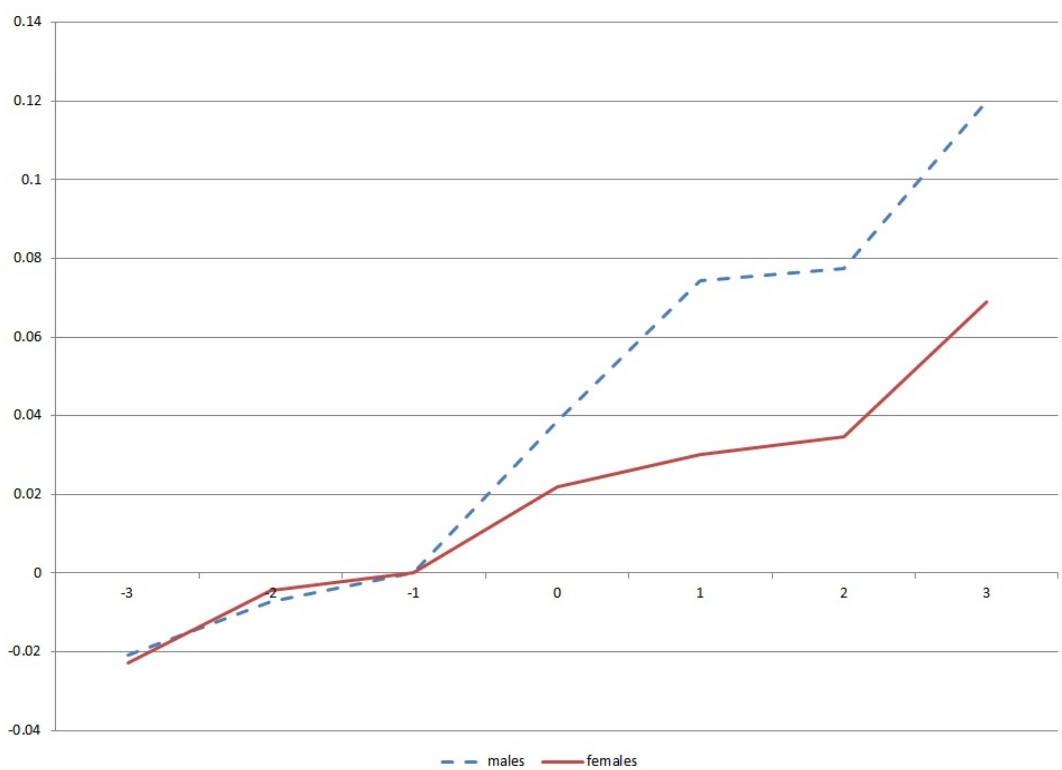


Figure 3.9: Signaling value of graduating from a university measured in log wages and net scores and intake effects on human capital.



Notes: Year 0 on the x-axis represents (aft0) the first cohort that the newly upgraded university took in. Year 1 (aft1) is the 2nd cohort. The base year is Year -1 (bef1), and so on.

Tables

Table 3.1: Summary Statistics for Data

A. Birth Data

	Freq.	Parity	Low	Age-mom	Age-dad	Income
1978	401,953	2.23	0.09	25.29	29.49	28,436
1979	410,947	2.16	0.08	25.29	29.28	29,158
1980	400,513	2.15	0.07	25.39	29.24	29,547
1981	404,730	2.11	0.07	25.50	29.21	30,372
1982	390,705	2.07	0.07	25.61	29.20	30,713
1983	370,622	2.05	0.07	25.75	29.27	31,085

B. College Data

	Freq.	Male	Urban	College	Score
1996	29,721	0.49	0.16	0.16	325.83
1997	61,339	0.52	0.17	0.17	360.91
1998	69,872	0.53	0.17	0.18	347.63
1999	71,952	0.53	0.16	0.18	332.39
2000	76,133	0.52	0.16	0.17	313.45
2001	72,037	0.52	0.16	0.09	345.38

C. First-Year Monthly Earnings Data

	Wage Year					
	2000	2001	2002	2003	2004	2005
1996	26,713	27,270	29,100	31,785	33,393	35,041
	(7310)	(6973)	(9549)	(11408)	(14162)	(15237)
1997		24,426	24,283	28,239	30,364	32,589
		(12793)	(12633)	(19890)	(26993)	(30508)
1998			22,166	25,737	28,354	30,821
			(11763)	(14214)	(25668)	(30972)
1999				22,188	25,468	28,353
				(10535)	(17078)	(25862)
2000					22,889	25,956
					(14935)	(19527)
2001						22,257
						(15446)

Notes: In Panel A, far-left column years represent years of birth. In Panel B, far-left column years represent years in which the student took the college entrance exam. In Panel C, far-left column years represent years of earnings; sample size for each cohort-by-wage-year subsample is in paranthesis.

Table 3.2: RD estimates of the first-stage outcomes.

	A. Male		
	Pr(HEI)	Pr(aft97|HEI)	Pr(Univ | HEI)
	(1)	(2)	(3)
1(Treat)	0.0482***	0.4625***	0.0946***
	(0.012)	(0.062)	(0.011)
N	1233303	199443	199443
Clusters	72	72	72
	B. Female		
	Pr(HEI)	Pr(aft97|HEI)	Pr(Univ | HEI)
	(4)	(5)	(6)
1(Treat)	0.0565***	0.5378***	0.0823***
	(0.011)	(0.068)	(0.014)
N	1146163	181607	181607
Clusters	72	72	72

Notes: Pr(HEI) is the probability of attending an HEI; Pr(aft97|HEI) is the probability of taking the JCEE after 1997, given that one is in the sample of college-goers. Pr(Univ|HEI) is the probability that you were enrolled at a university instead of college, given that one is attending an HEI. Standard errors clustered at the year-month of birth level. + $p<0.10$ * $p<0.05$ *** $p<0.01$

Table 3.3: DID estimates of the signaling value of university over college.

	A. H-FE		B. HD-FE		C. HD-FE 1997-2000	
	(1)	(2)	(3)	(4)	(5)	(6)
	m	f	m	f	m	f
UNIEVER	0.0616***	0.0423***	0.0430***	0.0487***	0.0445***	0.0290***
	(0.014)	(0.010)	(0.010)	(0.007)	(0.015)	(0.011)
UNIEVER_center	0.0190***	0.0221***	0.0084	0.0202***	0.0192+	0.0133*
	(0.006)	(0.004)	(0.005)	(0.003)	(0.011)	(0.007)
N	53912	88629	53912	88629	13049	24117
Clusters	5440	5985	5440	5985	1229	1304

	D. H-FE		E. HD-FE		F. HD-FE 1997-2000	
	(7)	(8)	(9)	(10)	(11)	(12)
	m	f	m	f	m	f
UPGRADE	0.0478***	0.0294***	0.0335***	0.0348***	0.0566***	0.0286*
	(0.013)	(0.009)	(0.009)	(0.007)	(0.017)	(0.011)
UPGRADE_center	0.0082	0.0149***	0.0024	0.0137***	0.0103	0.0101
	(0.006)	(0.004)	(0.005)	(0.003)	(0.012)	(0.007)
N	53743	88083	53743	88083	13049	24117
Clusters	5399	5927	5399	5927	1229	1304

Notes: The treatment variables are UNIEVER (your HEI has "University" in its name) and upgrade (taking JCEE after your upgraded). H-FE uses a fixed-effect of HEI only, HD-FE uses fixed effect for HD, and HD-FE 1997-2000 uses fixed effect for HD on only the sample of students in HEIs that upgraded sometime within 1997 to 2000. M=Male, F=Female. Standard errors clustered at the HEI-Dept-Year level. + $p<0.10$ * $p<0.05$ *** $p<0.01$

Table 3.4: Event Study Estimates of Signaling Value of University over College by Cohort.

dep_var=log-wage	male	female
	(1)	(2)
switch_bef3	-0.0618*	-0.0388*
	(0.0298)	(0.0165)
switch_bef2	-0.0548+	-0.0307+
	(0.0326)	(0.0162)
switch_yr	0.0376	0.0285*
	(0.0232)	(0.0133)
switch_aft1	0.1100***	0.0512***
	(0.0304)	(0.0161)
switch_aft2	0.1602***	0.0687***
	(0.0453)	(0.0226)
switch_aft3	0.1152+	0.1108***
	(0.0629)	(0.0312)
R-squared	0.333	0.186
N	15224	28223
Clusters	1251	1320

Notes: Switch_yr is first cohort of students after school upgrades. Switch_aft1 is the 2nd cohort. The base year is switch_bef1. M=Male, F=Female. Standard errors clustered at the HD-year level. + p<0.10 * p<0.05 *** p<0.01

Table 3.5: Event Study estimates on changes in scores, intake size, and log-wage before and after colleges that upgraded.

	A. Score		B. Intake		C. Log-wage	
	m	f	m	f	m	f
	(1)	(2)	(3)	(4)	(5)	(6)
switch_bef3	5.4294+	8.3514*	-1.6232	-0.2353	-0.0002	0.0001
	(2.8231)	(3.9717)	(2.969)	(2.888)	(0.000)	(0.000)
switch_bef2	0.9330	3.8380	-4.1319+	0.4048	-0.0209	-0.0228+
	(2.8546)	(3.9875)	(2.437)	(2.686)	(0.020)	(0.013)
switch_yr	6.0816+	9.6318*	-5.9043+	-10.6448***	-0.0072	-0.0045
	(3.2995)	(3.9424)	(3.047)	(3.493)	(0.017)	(0.010)
switch_aft1	20.8044***	23.7970***	-4.6503	-11.8563***	0.0387+	0.0220+
	(4.0386)	(5.0670)	(3.708)	(4.492)	(0.023)	(0.013)
switch_aft2	18.9286***	22.6756***	-15.0163***	-23.0389***	0.0743***	0.0302*
	(7.2136)	(8.4653)	(5.466)	(5.959)	(0.025)	(0.014)
switch_aft3	26.3756***	22.8895*	-15.6247*	-24.6749***	0.0776*	0.0348
	(8.8901)	(10.7877)	(7.053)	(8.193)	(0.038)	(0.021)
score					yes	yes
intake					yes	yes
R-squared	0.884	0.816	0.815	0.807	0.816	0.186
N	15224	28223	19428	31766	28223	28223
Clusters	1251	1320	1477	1513	1320	1320

Notes: Switch_yr is first cohort of students after school upgrades. Switch_aft1 is the 2nd cohort. The base year is switch_bef1. M=Male, F=Female. Standard errors clustered at the HD-year level. + p<0.10 * p<0.05 *** p<0.01

Table 3.6: Event Panel A: DID re-estimates of Signaling Value of university, net human capital effects and using UNIEVER.

	A. Male				B. Female			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
UNIEVER	0.0501***	0.0484***	0.0495***	0.0442***	0.0468***	0.0479***	0.0454***	0.0322***
	(0.009)	(0.009)	(0.009)	(0.015)	(0.008)	(0.008)	(0.008)	(0.011)
UNIEVER_center	0.0125***	0.0121***	0.0123***	0.0170+	0.0244***	0.0239***	0.0242***	0.0059
	(0.005)	(0.005)	(0.005)	(0.009)	(0.003)	(0.003)	(0.003)	(0.006)
FE	yes	yes	yes	yes	yes	yes	yes	yes
Controls	yes	yes	yes	yes	yes	yes	yes	yes
score	yes		yes	yes	yes		yes	yes
intake		yes	yes	yes		yes	yes	yes
1997-2000				yes				yes
R-squared	0.512	0.512	0.512	0.505	0.362	0.362	0.362	0.340
N	59860	59860	59860	16634	92005	92005	92005	27099
Clusters	5680	5680	5680	1455	6064	6064	6064	1493

Notes: Standard errors clustered at the HD-year level. + p<0.10 * p<0.05 *** p<0.01

Bibliography

Abdulkadiroğlu, Atila and Tayfun Sönmez, "School Choice: A Mechanism Design Approach," *American Economic Review*, Jun 2003, *93* (3), 729–747.

_ **, Josh D. Angrist, and Parag A. Pathak**, "The Elite Illusion: Achievement Effects at Boston and New York Exam Schools," Jul 2011.

Aghion, Philippe, Mathias Dewatripont, Caroline Hoxby, Andreu Mas-Colell, and André Sapir, "The Governance and Performance of Universities: Evidence from Europe and the US," *Economic Policy*, 2010, *25* (61), 7–59.

Altbach, P.G., L. Reisberg, and L.E. Rumbley, *Trends in global higher education: Tracking an academic revolution. A Report for UNESCO 2009 World Conference on Higher Education* 2009.

Angrist, Joshua D. and Jörn-Steffen Pischke, *Mostly Harmless Econometrics: An Empiricist's Companion*, Princeton University Press, 2009.

_ **and Jörn-Steffen Pischke**, "The Credibility Revolution in Empirical Economics: How Better Research Design is Taking the Con out of Econometrics," *Journal of Economic Perspectives*, Spring 2010, pp. 3–30.

Arcidiacono, Peter, "Ability Sorting and the Returns to College Major," *Journal of Econometrics*, Jul 2004, *121* (1-2), 343–375.

_ **, Patrick Bayer, and Aurel Hizmo**, "Beyond Signaling and Human Capital: Education and the Revelation of Ability," *American Economic Journal: Applied Economics*, 2010, *2* (4), 76–104.

Avery, Christopher and Jonathan Levin, "Early Admissions at Selective Colleges," *American Economic Review*, December 2010, *100* (5), 2125–56.

_ **, Andrew Fairbanks, and Richard Zeckhauser**, *The Early Admissions Game: Joining the Elite*, Harvard University Press, 2004.

Barreca, Alan I., Melanie Guldi, Jason M. Lindo, and Glen R. Waddell, "Saving Babies? Revisiting the Effect of Very Low Birth Weight Classification," *Quarterly Journal of Economics*, Nov 2011, *126* (4), 2117–2123.

Becker, Gary S., William H. J. Hubbard, and Kevin M. Murphy, "Explaining the Worldwide Boom in Higher Education of Women," *Journal of Human Capital*, 2010, *4* (3), pp. 203–241.

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan, "How Much Should We Trust Differences-In-Differences Estimates?," *Quarterly Journal of Economics*, 2004, *119* (1), 249–275.

Black, Dan A. and Jeffrey A. Smith, "Estimating the Returns to College Quality with Multiple Proxies for Quality," *Journal of Labor Economics*, 2006, *24* (3).

Bound, John, Brad Hershbein, and Bridget Terry Long, "Playing the Admissions Game: Student Reactions to Increasing College Competition," *Journal of Economic Perspectives*, 2009, *23* (4), 119–146.

Buser, Thomas, Muriel Niederle, and Hessel Oosterbeek, "Gender, Competitiveness and Career Choices," Technical Report, National Bureau of Economic Research 2012.

Carrell, Scott E., Bruce I. Sacerdote, and James E. West, "From Natural Variation to Optimal Policy? The Lucas Critique Meets Peer Effects," *National Bureau of Economic Research Working Paper Series*, 2011, *No. 16865.*

Chevalier, Arnaud and Gavan Conlon, "Does It Pay to Attend a Prestigious University?," *Institute for the Study of Labor (IZA) Working Paper*, 2003, *No. 848.*

Chou, Shin-Yi, Jin-Tan Liu, and James K. Hammitt, "National Health Insurance and Precautionary Saving: Evidence from Taiwan," *Journal of Public Economics*, Sep 2003, *87* (9-10), 1873–1894.

_ , _ , Michael Grossman, and Ted Joyce, "Parental Education and Child Health: Evidence from a Natural Experiment in Taiwan," *American Economic Journal: Applied Economics*, Jan 2010, *2* (1), 33–61.

Clark, Damon, "Selective Schools and Academic Achievement," *Institute for the Study of Labor (IZA) Working Paper*, 2007, *3182.*

Clark, Diana E. and Chang-Tai Hsieh, "Schooling and Labor Market Impact of the 1968 Nine-Year Education Program in Taiwan," Working Papers 215, Princeton University, Woodrow Wilson School of Public and International Affairs, Research Program in Development Studies. June 2000.

Currie, Janet, "Healthy, Wealthy, and Wise: Socioeconomic Status, Poor Health in Childhood, and Human Capital Development," *Journal of Economic Literature*, 2009, *47* (1), 87–122.

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

_ **and** _ , "Estimating the Return to College Selectivity over the Career Using Administrative Earnings Data," *National Bureau of Economic Research Working Paper Series w17159*, Jun 2011.

Dobkin, Carlos and Fernando Ferreira, "Do School Entry Laws Affect Educational Attainment and Labor Market Outcomes?," *Economics of Education Review*, 2010, *29* (1), 40–54.

Dong, Yingying, "Jumpy or Kinky? Regression Discontinuity without the Discontinuity," Technical Report, University Library of Munich, Germany 2010.

Duflo, Esther, Pascaline Dupas, and Michael Kremer, "Peer Effects, Teacher Incentives, and the Impact of Tracking: Evidence from a Randomized Evaluation in Kenya," *American Economic Review*, Aug 2011, *101* (5), 1739–1774.

Gindling, TH and Way Sun, "Higher education planning and the wages of workers with higher education in Taiwan," *Economics of Education Review*, 2002, *21* (2), 153–169.

Goastellec, Gaële, "Changes in Access to Higher Education: From Worldwide Constraints to Common Patterns of Reform?," May 2008, *International Perspectives on Education and Society* (9), 1–26.

Goldin, Claudia and Lawrence F. Katz, "Long-Run Changes in the Wage Structure: Narrowing, Widening, Polarizing," *Brookings Papers on Economic Activity*, 2009, (2), 135–165.

Hahn, Jinyong, Petra Todd, and Wilbert Van der Klaauw, "Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design," *Econometrica*, Jan 2001, *69* (1), 201–209.

Hanushek, Eric A. and Ludger Woessmann, "Do Better Schools Lead to More Growth? Cognitive Skills, Economic Outcomes, and Causation," *Journal of Economic Growth*, Jul 2012, pp. 1–55.

Hoekstra, Mark, "The Effect of Attending the Flagship State University on Earnings: A Discontinuity-Based Approach," *Review of Economics and Statistics*, 2009, *91* (4), 717–724.

Imbens, Guido and Karthik Kalyanaraman, "Optimal Bandwidth Choice for the Regression Discontinuity Estimator," *National Bureau of Economic Research*, Feb 2009, (No. 14726).

Imbens, Guido W, "Better LATE Than Nothing: Some Comments on Deaton (2009) and Heckman and Urzua (2009)," *Journal of Economic Literature*, Jun 2010, *48* (2), 399–423.

Imbens, Guido W. and Thomas Lemieux, "Regression Discontinuity Designs: A Guide to Practice," *Journal of Econometrics*, Feb 2008, *142* (2), 615–635.

Jackson, C. Kirabo, "Do Students Benefit from Attending Better Schools? Evidence from Rule-based Student Assignments in Trinidad and Tobago," *The Economic Journal*, Dec 2010, *120* (549), 1399–1429.

Jacob, Robin, Pei Zhu, Marie-Andrée Somers, and Howard Bloom, "A Practical Guide to Regression Discontinuity," *MDRC Report #664*, July 2012.

Lee, David S and Thomas Lemieux, "Regression Discontinuity Designs in Economics," *Journal of Economic Literature*, Jun 2010, *48* (2), 281–355.

Liu, Jin-Tan, Shin-Yi Chou, and Jin-Long Liu, "Asymmetries in progression in higher education in Taiwan: Parental education and income effects," *Economics of Education Review*, 2006, *25* (6), 647–658.

Lucas, Robert E., "On the Mechanics of Economic Development," *Journal of Monetary Economics*, July 1988, *22* (1), 3–42.

MacLeod, W. Bentley and Miguel Urquiola, "Anti-Lemons: School Reputation, Relative Diversity, and Educational Quality," 2011.

McCrary, Justin, "Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test," *Journal of Econometrics*, 2008, *142* (2), 698–714.

_ **and Heather Royer**, "The Effect of Female Education on Fertility and Infant Health: Evidence from School Entry Policies Using Exact Date of Birth," *American Economic Review*, February 2011, *101* (1).

Niederle, Muriel and Lise Vesterlund, "Do Women Shy Away from Competition? Do Men Compete too much?," *Quarterly Journal of Economics*, 2007, *122* (3), 1067–1101.

Oreopoulos, Philip, "Estimating Average and Local Average Treatment Effects of Education when Compulsory Schooling Laws Really Matter," *American Economic Review*, March 2006, *96* (1).

_ **and Kjell G. Salvanes**, "Priceless: The Nonpecuniary Benefits of Schooling," *Journal of Economic Perspectives*, 2011, *25* (1), 159–184.

_ **, Till von Wachter, and Andrew Heisz**, "The Short- and Long-Term Career Effects of Graduating in a Recession," *American Economic Journal: Applied Economics*, Jan 2012, *4* (1), 1–29.

Pop-Eleches, Cristian and Miguel Urquiola, "Going to a Better School: Effects and Behavioral Responses," *American Economic Review*, 2012.

Rubin, Donald B., "Causal Inference Using Potential Outcomes," *Journal of the American Statistical Association*, 2005, *100* (469), 322–331.

Rubinstein, Yona and Sheetal Sekhri, "Do Public Colleges in Developing Countries Provide Better Education than Private ones? Evidence from General Education Sector in India," *Centre for the Economics of Education Discussion Paper #0130*, 2011.

Saavedra, Juan E., "The Returns to College Quality: A Regression Discontinuity Analysis," 2008, (Unpublished Manuscript).

Spence, Michael, "Job Market Signaling," *Quarterly Journal of Economics*, 1973, *87* (3), 355–374.

Spohr, Chris A., "Formal Schooling and Workforce Participation in a Rapidly Developing Economy: Evidence from "Compulsory" Junior High School in Taiwan," *Journal of Development Economics*, Apr 2003, *70* (2), 291–327.

Tao, Hung-Lin, "Monetizing college reputation: The case of Taiwan's engineering and medical schools," *Economics of Education Review*, 2007, *26* (2), 232–243.

Tsai, Shu-Ling and Yossi Shavit, "Taiwan: Higher Education–Expansion and Equality of Educational Opportunity," in Y. Shavit, R. Arum, A. Gamoran, and G. Menachem, eds., *Stratification in higher education: A comparative study*, Stanford Univ Pr, 2007, pp. 140–164.

Tsai, Wehn-Jyuan, Jin-Tan Liu, Shin-Yi Chou, and Michael Grossman, "Intergeneration Transfer of Human Capital: Results from a Natural Experiment in Taiwan," *National Bureau of Economic Research Working Paper Series w16876*, 2011.

_ , _ , _ , and Robert Thornton, "Does Educational Expansion Encourage Female Workforce Participation? A Study of the 1968 Reform in Taiwan," *Economics of Education Review*, 2009, *28* (6), 750 – 758.

Tyler, John H., Richard J. Murnane, and John B. Willett, "Estimating the labor market signaling value of the GED," *Quarterly Journal of Economics*, 2000, *115* (2), 431–468.

Urquiola, Miguel and Eric Verhoogen, "Class-Size Caps, Sorting, and The Regression-Discontinuity Design," *American Economic Review*, 2009, *99* (1), 179–215.

Van der Klaauw, Wilbert, "Estimating the Effect of Financial Aid Offers on College Enrollment: A Regression–Discontinuity Approach," *International Economic Review*, 2002, *43* (4), 1249–1287.

Vere, James P., "Education, Development, and Wage Inequality: The Case of Taiwan," *Economic Development and Cultural Change*, April 2005, *53* (3), 711–735.

Vigdor, Jacob L. and Charles T. Clotfelter, "Retaking the SAT," *Journal of Human Resources*, 2003, *38* (1), 1–33.

Weiss, Andrew, "Human Capital vs. Signalling Explanations of Wages," *Journal of Economic Perspectives*, 1995, *9* (4), 133–154.

Winston, Gordon C., "Subsidies, Hierarchy and Peers: The Awkward Economics of Higher Education," *Journal of Economic Perspectives*, Jan 1999, *13* (1), 13–36.

Wiske-Dillon, Eleanor and Jeffrey Smith, "Determinants of Mismatch between Students and Colleges," 2012, (Unpublished Manuscript).

Yang, Chih-Hai, Chun-Hung A. Lin, and Chien-Ru Lin, "Dynamics of Rate of Returns for Postgraduate Education in Taiwan: The Impact of Higher Education Expansion," *Asia Pacific Education Review*, Nov 2010, *12* (3), 359–371.