Cross-Program Differences in Returns to Education and the Gender Earnings Gap

Steffen Andersen, Philippe d'Astous, Jimmy Martínez-Correa and Stephen H. Shore*

June 3, 2020

Abstract

University programs differ in their gender earnings gaps, the difference between the subsequent earnings of the program's male and female enrollees. A program could have a positive gender earnings gap because the program attracts higher-ability men than women (a selection effect), or because the program increases the earnings of male enrollees more than female enrollees (a causal effect). To understand the source of cross-program differences in gender earnings gaps, we exploit a discontinuity built into the Danish national university admissions system, which provides a quasi-random assignment of similar applicants to different programs. Enrolling in a program with a $1 larger gender earnings gap, holding average earnings constant, does not affect male earnings but reduces female earnings by $0.45. This effect is small as women enter the labor market but increase over time, and cannot be explained by channels related to marriage or childbirth. Our results show that programs that *appear* worse for women – in the sense of having economically significant gender earnings gaps – *are* worse for women because they reduce female earnings more than programs with smaller gaps.

JEL Classification: I23, I26, J24

*Andersen: Department of Finance, Copenhagen Business School, Solbjerg Plads 3, DK-2000 Frederiksberg, Denmark, email: san.fi@cbs.dk. d'Astous: Department of Finance, HEC Montréal, 3000 Chemin de la Côte-Sainte-Catherine, Montréal, Québec, H3T 2A7, Canada, email: philippe.dastous@hec.ca. Martínez-Correa: Department of Economics, Copenhagen Business School, Porcelaenshaven 16A, DK-2000, Frederiksberg, Denmark, email: jima.eco@cbs.dk. Shore: Department of Risk Management and Insurance, Robinson College of Business, Georgia State University, 35 Broad Street, Atlanta, GA 30303, USA, email: sshore@gsu.edu. We thank the Economic Policy Research Network (EPRN) housed at the University of Copenhagen, the Pension Research Center (PeRCent) at the Copenhagen Business School, the Retirement and Savings Institute at HEC Montréal, the European Research Council (ERC), and the Center for the Economic Analysis of Risk at Georgia State University for financial support. We thank Matias Cattaneo, Alejandro Del Valle, Fane Groes, Dan Kriesman, Seth D. Zimmerman and seminar participants at the 2020 American Economics Meetings, 2019 CEAR Registry Data and Experiments Workshop, Georgia State University, and the Vienna Graduate School of Finance for many useful comments, as well as Ann-Lea Falk Larsen for excellent and dedicated research assistance.

1 Introduction

Some university programs appear relatively worse than others for women in the sense that they have larger gender earnings gaps; their male enrollees go on to earn more than their female enrollees. For example, programs within the broad field of social science differ substantially in the subsequent gender earnings gaps of their enrollees, with much larger gaps in economics programs than in psychology programs.[1] Even among economics programs, there are substantial differences between universities in terms of their programs' gender earnings gaps. Large gender earnings gaps could reflect selection effects (attracting higher-ability men than women) or causal effects (increasing earnings for men more than for women). What are the relative roles of selection and causal effects in explaining the cross-program differences in the gender earnings gap?

To answer this question, we exploit an institutional feature of the Danish university admissions system that quasi-randomly assigns university applicants to different programs. Prospective university students receive a composite score summarizing their performance in high school. They then rank their preferred programs (university–major pairs, such as physics at the University of Copenhagen) from first to the eighth choice, at most. Selective programs have a score cutoff, which changes from year to year to clear the market. Students are admitted to their most preferred program for which their score meets or exceeds the cutoff.

We focus on students who have a score that is close to the cutoff for their most preferred "reach" program and well above the cutoff for their next most preferred "safety" program. For students in this "reach-safety sample", a small increase in their score from below to just above the cutoff of their reach program leads to a large and discontinuous increase (decrease) in the probability of enrolling in their reach (safety) program. Our identification strategy exploits this quasi-random assignment of similar students to programs with varying gender earnings gaps in a fuzzy regression discontinuity IV design. Consider a student whose reach program has a larger gender earnings gap than their safety program but the two programs have the same average earnings.[2] If cross-program differences in the gender earnings gap reflect causal effects, crossing

[1]For example, in Denmark, the annual earnings gaps are 56,000 DKK for economics and 23,000 DKK for psychology. The difference between economics and psychology programs in terms of their enrollees' subsequent average earnings is proportionally much smaller (380,000 DKK and 280,000 DKK, respectively).

[2]For example, imagine male enrollees in the reach program subsequently earn 450,000 DKK on average while female enrollees subsequently earn 350,000 DKK on average; men and women are equally represented in the reach program, so that average earnings for the program are 400,000 DKK overall. Male and female enrollees in the safety program

this threshold from below – so the student becomes discontinuously more likely to enroll in the higher-gap reach program – will increase earnings for male students and decrease earnings for female students.

We test this hypothesis and find that enrolling in a program with a larger gender earnings gap, holding average enrollee earnings constant, reduces the subsequent earnings of women but not men. We estimate that enrolling in a program whose enrollees experience a $1 larger gender earnings gap, controlling for earnings, leads to women earning $0.45 less on average (C.I. –$0.81 to –$0.096) when evaluated at our baseline time horizon of 10 to 12 years after enrollment.[3] This result is also statistically indistinguishable from OLS estimates, which suggests that cross-program differences in the gender earnings gap reflect cross-program differences in the relative returns experienced by men and women, and not a differential selection by men and women into various programs. In other words, we would expect that if a woman were to enter a program with a 100,000 DKK larger gender earnings gap but the same average earnings, we would expect her annual earnings to be 45,000 DKK lower.

The causal effect for women of their program's gender earnings gap is marginal when women initially enter the labor market seven to nine years after enrollment and grows over time to $0.80 (C.I. –$1.24 to –$0.36) 13 to 15 years after enrollment. Such result is in line with previous research findings that the observed gender earnings gap tends to be low as graduates enter the labor market, and then increases over time (Bertrand et al., 2010, Goldin et al., 2017).

We also show that the causal effect of average program earnings is stronger for women than for men. Women earn more when they enter programs with higher average earnings, but men do not. We estimate that women earn an additional $0.48 by enrolling in a program with an additional $1 of average program earnings (C.I. $0.23 to $0.74) while men earn an (insignificant) additional $0.06 (C.I. –$0.31 to $0.43). These IV estimates are almost statistically indistinguishable from the OLS results for women ($0.79, C.I. $0.73 to $0.85) but are very different for men ($0.75, C.I. $0.65 to $0.85). The similarity of IV and OLS estimates for women suggest that the causal effect of program earnings is the key driver of cross-program differences in average program earnings for

both earn 400,000 DKK on average. The gender earnings gap is 100,000 DKK for the reach program and zero for the safety program.

[3] In our baseline 10 to 12 years horizon, individuals are in their early thirties and they have been approximately 5 years in the labor market.

women; the divergence of IV and OLS estimates for men suggests that selection is the key driver of cross-program differences in program earnings for men. The earnings of women depend much more on the average earnings of the program they enter than do the earnings of men.[4]

Our initial analysis provides evidence that enrolling in programs with large gender earnings gaps reduces earnings for women, but does not tell us why. First, we look for possible pathways, namely factors that may impact earnings and be impacted by program. For example, marriage and the presence of children are associated with lower earnings for Danish women; some programs may reducing earnings for women by increasing the prevalence of marriage and children. To explore this hypothesis, we add marriage and child controls to our regressions, but find that our results persist. This suggests that marriage and children are not channels through which programs' gender earnings gaps operate to reduce earnings for women. Second, we examine the level – academic institution, broad field of study, or specific program – at which the presence of a gender earnings gaps operates to reduce earnings for women. We find that the impact of gender earnings gaps on earnings for women are similar whether those gaps are measured at the level of the institution, field, or program.

A few studies have used admissions discontinuities similar to the Danish ones we exploit to study different questions: Kirkeboen et al. (2016) in Norway; Hastings et al. (2013) in Chile; and Heinesen (2018), Humlum et al. (2017), Foley and Groes (2018), Daly and Le Maire (2019) and Heinesen and Hvid (2019) in Denmark. Previous studies have used such discontinuities to identify the effect of entering specific programs of study on average earnings. Unlike those studies, we examine cross-program differences in the gender earnings gap to test how much of the dispersion

[4]Our results can be further explained with a hypothetical example of a female student who considers enrolling in the Finance program (a business field) at the University of Århus or to the Psychology program (a Social Sciences field) at the Copenhagen University. Imagine that there are equal numbers of men and women in both programs; the average annual earnings in the Psychology program is 450,000 DKK, and there is no gender gap in this program so the average incomes for men and women are equal; furthermore, the average earnings in this Finance program are 500,000 DKK but there is a 100,000 DKK gender gap, with women earning 450,000 DKK and men earning 550,000 DKK on average. We can interpret our point estimate on average earnings for women of 0.48 as the gain for a female student who would choose Finance instead of Psychology. In this case, this represents an increase in average earnings of 24,000 DKK (a coefficient of 0.48 multiplied by 50,000 DKK, the difference between the programs in average earnings).

Our gender gap point estimate for women of 0.45 represents the cost for women of enrolling in a program with a larger gender earnings gap. In this case, it represents a 45,000 DKK penalty for a female student studying the Finance program instead of the Psychology program (a coefficient of 0.45 multiplied by 100,000 DKK, the difference between the programs in average gender earnings gap). In particular, in this example a female student would earn 21,000 DKK more (45,000 DKK minus 24,000 DKK) by enrolling in the Psychology program instead of the Finance program, even though Finance has higher average earnings.

reflects selection versus causal effects.[5]

One challenge inherent in estimating the return from a program of study is that there are many programs, and consequently many potential parameters to identify (e.g. the causal effect of moving from program X to program Y). The number of university–major pairs available in any given year is large, and the number of possible combinations of programs for which one might compute causal effects is enormous.[6] Instead of focusing on specific program pairs, we describe each program in terms of the gender earnings gap of its enrollees, and measure the effect on earnings of men and women of moving across programs with small to large gender earnings gap. Our identification strategy shrinks the problem and allows us to estimate differences in returns experienced by men and women instead of estimating pair-wise returns *between* various broad fields. This approach has the appealing feature that it maps to the actual signal extraction problem faced by many prospective university students who observe the earnings of various programs and wish to estimate how entering a particular program will affect their own earnings.

Another potential challenge for our empirical strategy is that students may try to use the information about their scores and historical cutoffs of their preferred programs strategically. While honest ranking of preferences is optimal in the Danish admissions system,[7] if prospective students can guess the admission cutoffs, students who do not expect to enroll in a program may be less likely to rank it. Such bunching could present an econometric problem if the likelihood of bunching was correlated with unobserved ability. We perform various manipulation checks to rule out this possibility and ensure that the institutional setting generates exogenous variation in student scores relative to the cutoff score of students' preferred reach programs. We present density tests following Calonico et al. (2018) and reject any manipulation of the assignment variable. We further

[5]Heckman et al. (2018) discusses two general approaches to estimate the causal effects of education: structural models (e.g. Keane and Wolpin (1997) and Eisenhauer et al. (2015)) and instrumental variables methods including both randomization and regression discontinuities as instruments (e.g. D.Angrist and Imbens (1995)).

[6]Kirkeboen et al. (2016) and Heinesen and Hvid (2019) deal with this problem by estimating a series of pair-wise comparisons across broad fields of study or universities, effectively measuring the effect of entering one broad category of programs over another (e.g. studying in a health-related program versus one that is engineering-related, or a program from one university versus that of another). Using a similar approach, Hastings et al. (2013), Humlum et al. (2017), Foley and Groes (2018) and Daly and Le Maire (2019) classify educational programs into broad fields. They estimate the effect of entering the most preferred educational program as opposed to a mix of next-best degrees. Heinesen (2018) estimates the causal effects of being accepted into the first-ranked program as opposed to a program that is not the first choice on earnings and educational completion.

[7]The admission system is similar to the Swedish and the Norwegian systems and are all closely related to serial dictatorship, subject to limits in the number of programs that can be ranked, which is both Pareto efficient and strategy-proof (Svensson, 1999).

present evidence that cutoffs change from year to year. We use this fact in the analysis to obtain clean identification; while students can know last year's cutoffs, they do not know the cutoff they will face when applying. As a result, by controlling for a student's score relative to last year's cutoffs, our estimated parameter is identified by the change in probability of entering a program, given year-to-year changes in program cutoffs. Finally, we present evidence of covariate balance around the admission cutoffs, suggesting that students on each side of the cutoffs are comparable on observable characteristics.

The literature on estimating the return to education is large, with an extensive focus on estimating the impact of additional years of schooling on various outcomes.[8] Fewer studies examine the causal effect of admission into different programs of study.[9] Research on the gender earnings gap is also substantial.[10] A large part of the literature documents the unconditional gap in earnings between men and women, and decomposes that gap into observable factors (for example, experience and chosen occupation) and unobservable factors (for example, discrimination, as in Blau and Kahn (2017)). We take a different approach by comparing university program differences in the gender earnings gap instead of trying to explain the overall gap.In this regard, we know of one other study, (Hausmann et al., 2017), which examines differences across industries and occupations in terms of field-specific female representation. Of other evidence, the National Science Foundation (2017) reports that women receive roughly 25% of all PhD degrees awarded in mathematics and statistics.[11] Stark gender earnings gaps in particular fields have recently been documented in news reports (Cain and Kiersz, 2018). There has been a recent focus on the particular challenges and experiences of women within the economics profession. For example, National Science Foundation (2017) and Lundberg (2018) have noted the under-representation of women at all levels of the profession. Wolfers (2018) also notes that "the list of problems [for women in

[8]For effects on earnings, see literature reviews from Psacharopoulos and Patrinos (2018) and Tables 4, 5 and 6 in Card (1999). See Lochner (2011) for a literature review on the effect on criminality, health, mortality, citizenship and political participation. See Hanushek and Rivkin (2006) for a literature review on the effect of teacher quality, class size and teacher demographics. Research from Hoxby (2009) and Dale and Krueger (2002) estimates the impact of attending one university over another.

[9]Altonji et al. (2012) and Altonji et al. (2016) review the literature on payoffs in relation to field of study or major. Arcidiacono (2004) handles selection with a dynamic discrete choice model, and Wiswall and Zafar (2015) use an information experiment wherein they measure how beliefs change after providing objective statistics of earnings by major.

[10]For example, see Ñopo et al. (2011) for a review of evidence around the world. See also Bishu and Alkadry (2017) for a systematic review of gender pay gap evidence in both the economics and non-economics literature.

[11]Bertrand et al. (2010) studied gender disparities in the corporate and finance sectors. Reuben and Wiswall (2016) studied gender differences in terms of major choice.

economics] is daunting."[12]

The results we document in this study have important implications for policy and our understanding of the gender earnings gap. Our results suggest that interventions in programs with large gender earnings gaps, or the fields they feed into, may be useful in reducing gender earnings gaps and increasing female earnings in those programs. Second, the fact that the gender earnings gap is present both at the university and field levels suggests that educational institutions have a role to play in preventing discrimination against their students, whether while they are in school or once they reach the job market. Lastly, because we find that the gender earnings gap grows over time, more results are needed to document the career progression of women in their jobs to identify the extent and mechanism of gender discrimination.

2 Institutional Setting and Data

2.1 Institutional Setting

Admissions to Danish universities are handled centrally by a government agency called *Den Koordinerede Tilmelding*, which translates to the Coordinated Signup. Admission generally requires a high school degree, from which students receive grades for each course they completed in upper secondary school. An algorithm calculates a standardized score, varying in 0.1 increments from 0 to 13, as the weighted average of their course grades, with weights that vary based on the difficulty of the courses (e.g. A-level courses typically receive greater weights than B-Level courses, which in turn typically receive greater weights than C-level courses). The admission process is the same for a vast majority of programs. The high school score is often the only relevant factor considered for admission, although certain programs, such as medicine, may also have requirements on the course grade or the level of specific subjects.

When applying to university, prospective students can rank up to 8 different programs (i.e. university major pairs, such as economics at the University of Copenhagen) in preferred order.

[12]Wolfers (2018) notes that "women in the field are often held to higher standards in relation to their written work (Hengel, 2017) and are not given credit for papers written collaboratively with men (Sarsons, 2017). Student teaching evaluations tend to be biased against women (Boring et al., 2016), while journalistic discussions of economic research often relegate women to a secondary role (Wolfers, 2015)... In addition, a popular online discussion forum has been found all too often to sexualize or trivialize women and their work (Wu, 2018). Another study (Ceci et al., 2014) has found economics to be an 'outlier' among academic fields because of 'a persistent sex gap in promotion that cannot readily be explained by productivity differences'."

To be admitted into their preferred program, prospective students compete for a finite number of places. The Danish Ministry of Education and the universities, through their funding and priorities, determine the maximum intake of students in each educational program.[13] Places in each program are allocated to the students with the highest scores until all available places are filled. Programs therefore have admission cutoff scores that clear the market given the supply and demand for places and students are admitted to their most preferred program for which their score meets or exceeds the program's cutoff. The cutoffs from prior years are known when applying, but not the cutoffs relevant for the current application season.

While the key stylized features of the admission system are simple, some details are worth noting. First, two different admission systems, called "Quotas", allow students to apply to university. Students who come out of high school directly (without additional relevant experience) typically apply through "Quota 1" in June, and their admission depends on their high school score alone. Students with relevant post-high-school qualifications can apply through "Quota 2" in March, and a bonus may be added to their score to reflect the relevance of their additional qualifications to the program of study for which they are applying.[14] In this sense, Quota 1 admissions are the cleanest because they involve students without additional post-high-school qualifications and depend solely on the composite high school score that we observe. Quota 2 admissions are noisier because applicants can receive a bonus to their score that we do not observe based on the relevance of their extra-curricular qualifications. 54.4% of all applications are made through the Quota 1 system and 45.6% through the Quota 2 system. In our analysis, we retain applications made through both admission systems, but we treat them separately by adding Quota fixed effects and calculating their admission probabilities separately.

Second, students can choose to apply with the option of being on a standby list in case they are very close to the program's cutoff but are not initially admitted to the program (about 40% of students select this option). The standby list avoids leaving empty places in programs where

[13]The Ministry of Education regulates the numbers of students graduating within specific programs with the purpose of satisfying the predicted demand for given qualifications in the Danish labor market. As all Danish universities are public, there are no tuition fee constraints on poor students. Students also receive educational subsidy support (*Statens Uddannelsesstøtte* in Danish) and can apply for generous student loans. Another detail omitted from the description is the option for students to postpone enrolling in a higher education program.

[14]These could include completion of additional courses, work experience, or any other additional qualifications deemed relevant for a given program. We do not observe these additional points in our data. Quota 2 applications are initially evaluated within the Quota 1 system, and for the applications that do not make the cut on Quota 1, they are then evaluated in Quota 2 and the additional qualifications beyond the high school grades are taken into account.

admitted students either reject the offer or quickly drop out of the program. The cutoff for the standby list is usually 0.1 to 0.2 grade points below the actual cutoff point. If students choose to apply for standby places, they will be offered the standby place instead of their safety program if they are above the cutoff point for the standby list. In case they are not offered a place within the first month of the program starting, they are guaranteed a place in the program the following year. Because standby applications are conceptually different, in this study we focus on students who do not use this option. Nonetheless, we present the admission discontinuities for each admission system in Figure A1 and Section 5.5 we show that our results are robust to including the standby applications, and also hold for the Quota 1 subsample of applications.

Because the effects that we document are the results of infra-marginal variation in admissions cutoffs, caution is nevertheless warranted when extrapolating to students far from admission cutoffs. Furthermore, the assignment scheme produces variation at the admission-level, and not at the graduation-level. For this reason, we perform the analysis at the admission level and analyze program switching behavior in Section 5.1. We then re-estimate our baseline results using the program in which students are enrolled four years after enrollment and show similar results.

2.2 Data

We derive data on the entire Danish population from three different administrative sources made available through Statistics Denmark and detailed below. First, the Danish Civil Registration System provides demographic information including individuals' unique personal identification number, as well as their gender, date of birth, and marital history up to 2016. The personal identification number is used to merge information from the different administrative data sets. These administrative records also contain a unique household identification number, as well as the personal identification number of each individual's spouse and children in the household. We use these data to obtain demographic information about each individual and household, and to generate intergenerational links.

Second, the Ministry of Education provides the full list of programs to which prospective students applied in ranked order of preference between 1993 and 2006. It also includes upper secondary school grades, official cutoff information for each program, the program in which the individual has been accepted (if any), as well as the distinction between applications made through Quotas 1 and 2, and standby lists.[15] We use grade information from upper secondary education diplomas to identify the distance from a student's grade to the admission cutoff for a program. Although we observe the specific programs to which students apply and to which they are originally admitted, for subsequent attendance and graduation information, we only observe the institution code and a four-digit education code constructed by Statistics Denmark that identifies the field of study. We rely on this education code to follow what students are studying over the years after initial enrollment and the field from which they graduate.

Finally, the Danish Tax Authority provides disaggregated income and wealth information between 1993 and 2016. The Danish Tax Authority receives this information directly from relevant third-party sources: employers supply statements of wages paid to their employees, and all financial institutions supply information on their customers' information. Because taxation in Denmark mainly occurs at the source, this income and wealth information is highly reliable. For our purposes, the records include annual income and total net wealth. The earnings measures we use in the analysis represent labor income and exclude self-employment income and income earned from investments. We are able to link students to their parents and obtain parental income and wealth details. Although this information is extensive, not all components of wealth are recorded by the Danish Tax Authority: we do not have information about individuals' holdings of unbanked cash, their private debt (i.e. debt to private individuals), accumulated pension savings (although we observe pension contributions), private business equity, or other informal wealth holdings.

2.3 Sample Construction

2.3.1 Population data

We start by constructing the universe of students applying to a university program for the first time between 1994 to 2006. We restrict applications to these years for two reasons: (1) because

[15]All information is obtained by third parties: upper secondary grades from diplomas provided by the grading institutions, applications from the governmental agency collating all applications and clearing the market, and completed educational programs from the diploma granting institutions. Each program has a unique identifier and consists of a university–major pair, both of which are linked to the program identifier in the form of an institution code number and a major name. We verified that the unique program identifiers are associated with the same program every year in the sample we studied and dropped 15 programs for which the name changed substantially over time within a program identifier.

we look at outcomes for students at least 10 years after applying to university and most of our data are available until 2016; and (2) because we control for admission cutoffs in the year prior to application, and we have these data starting in 1993.

We only retain students for whom we have the full range of information used in our analyses, which gives a total of 320,308 student applications. We winsorize the financial outcomes at the 1% and 99% levels of this population of first-time applicants. We construct average earnings and gender earnings gaps across programs, fields of study and institutions, as well as segmented across men and women, allowing us to investigate the cross-sectional dispersion of various measures.

2.3.2 Reach-safety pairs

We then restrict our main analysis to the sample of prospective university students who are "close" to the discontinuity we exploit and have a viable pair of reach–safety programs of more and less preferred rankings. To be included in the reach-safety sample, students must have a reach program for which their score is within 0.5 of the admission cutoff (either above or below), such that students could plausibly be assigned to a different program if the admission cutoff happened to be slightly higher or lower that year.[16] Additionally, the student's next-most-preferred safety choice must be one for which the student's grades are easily good enough for them to be admitted (their score must be more than 0.5 above the admission cutoff). The restrictions that a student's score must be within 0.5 of the reach cutoff and 0.5 above that of the reach is tight in the sense that the standard deviation of scores is 1.0 in the overall sample of applicants.

10,944 prospective students have scores and program rankings that include them in the sample of applicants with viable reach–safety pairs, where small changes in student grades or reach cutoffs would plausibly yield assignment to a different program.

2.4 Admission Discontinuity

The key variable driving admission for student i is $D_i = S_i - C_{r_i}$, the difference between their score (S_i) and the cutoff for their reach program (C_{r_i}). When the assignment scheme is precisely implemented, students are admitted to a program (p) that is either their reach program (r) when

[16] We also test robustness to different bandwidths around the admission cutoff in Section 5.3.

$S_i \geq C_{r_i}$, or their safety program (s) otherwise:[17]

$$p = \begin{cases} r & \text{if } D_i \geq 0 \\ s & \text{if } D_i < 0. \end{cases} \tag{1}$$

However, as detailed above, the Danish university system contains complications that lead enrollment outcomes to deviate from the clean, deterministic predictions in equation (1). This is illustrated in Figure 1, which presents the share of individuals who are admitted to different programs as a function of D_i.[18] Panel (a) shows the subsample of applications made through the Quota 1 system of admission. The probability that students enroll in their reach program, $\mathbb{P}(p_i = r_i | D_i)$, jumps discontinuously from 10% to more than 90% as the difference between their score S_i and the cutoff of their reach program C_{r_i} crosses from negative to positive. Meanwhile, the probability that they enroll in their safety program, $\mathbb{P}(p_i = s_i | D_i)$, falls discontinuously. When necessary, admissions may be randomized among students with scores that fall exactly on the threshold, i.e. when $D_i = 0$. For this reason in Section 5.3 we also run a donut-hole analysis excluding these individuals. Panel (b) of Figure 1 shows admissions through the Quota 2 system, which are accordingly more noisy. In the analysis, we estimate a fuzzy instrumental RDD to take into account the fact that admission in each of these two systems is not a perfect function of the assignment variable. We measure time-invariant admission probabilities separately and include Quota fixed effects for each admission system.

Figure 1 shows that students also have a small probability of enrolling in other programs (neither reach nor safety programs, denoted by o_i) or in nothing (denoted by n_i). Some students elect to defer or not attend university, and some subsequently re-apply to switch programs. We con-

[17]For example, consider a hypothetical student who obtained a composite score of 9.8. The student does not have post-high-school qualifications, and thus applies under "Quota 1" based on this score alone. The student selects physics at the University of Copenhagen as the first choice, and chemistry at the University of Copenhagen as the second choice. We refer to the student's first and second choices (detailed in Section 2.3) as the "reach" and "safety" programs, respectively, because the student's grade of 9.8 turns out to be close to the 10.0 cutoff of the first choice, while it is much higher than the cutoff of 9.0 for admission to the second choice. In this example, the student is denied a place in the first/reach choice of physics because his/her grade falls just below its cutoff, and is admitted to the second/safety choice of chemistry instead. We could imagine that physics might have been less competitive the previous year, with a cutoff of only 9.7, in which case the student would have been admitted to study physics had the student applied the previous year.

[18]Cutoffs are not directly observed for all programs because of missing data. For approximately 13% of the programs, we impute the cutoff point. Mechanically, we classify a jump of more than 50 percentage points in the acceptance rate of Quota 1 students as the identified cutoff point. Excluding these programs does not change our conclusions.

sider these subsequent enrollment and switching issues in Section 5.1. However, because these outcomes are small and do not change drastically with D_i, in the empirical model we only exploit the variation in the acceptance to the reach and safety programs.

2.5 Summary Statistics of the Data

Table 1 presents descriptive statistics for the population of students applying to university for the first time in application years 1994 to 2006 (all applications; left) as well as the sample of students who have a reach–safety programs pair in these application years (used sample; right). It is helpful to contrast both samples to highlight the potential differences when evaluating the external validity of our study. We show that our reach-safety sample is somewhat positively selected on observables, which all else constant does not suggest that the effect of the gender earnings gap should be upward biased. Panel A presents demographics. In both the population and in the sample used, prospective students are on average 21 years old, and close to 39% are men. Parental income, wealth and education are the average values for both parents (or the values for the single parent on record) at the time of application. Students in the reach–safety programs pair sample have parents with higher incomes (328,600 DKK versus 298,100 DKK), higher wealth (432,900 DKK versus 348,700 DKK) and more education (13.5 years versus 12.7 years).[19] Panel B presents information typically found in a student's application. 60% of the sample are admitted to their reach program, which is on average listed as their first choice, while their safety program is on average listed as their second choice. On average, the program in which the students enroll is ranked at position 1.6. Panel C shows 60% of students have a reach–safety programs pair in the same broad field of study, and 30% have one at the same institution.

Panel D of Table 1 presents information on earnings. The earnings are averaged over the period 10 to 12 years after the applications to university programs, although we investigate earnings from 7 to 15 years after application in Section 4.3.[20] The average earnings of students in the reach–safety programs pair sample is similar to the overall population (306,900 DKK versus 300,400 DKK), and

[19] The January 2019 exchange rate for USD to DKK is 6.5.

[20] We maximize the number of years we can use in our sample by imposing the condition that the observation has at least 10 years of observations, and is averaged over future years 10 to 12 whenever possible. In particular, for the 2006 application-year cohort, we only have the outcomes 10 years after application, for the 2005 cohort we can average the outcomes 10 and 11 years after application, and for all prior years up to 2004 we can average the outcomes from 10 to 12 years after graduation. We do so to maximize the number of cohort years used in the analysis.

this is true for both women (286,600 versus 273,900) and men (338,600 versus 345,100). The students' enrollment cohort average earnings (303,600 DKK versus 294,100 DKK in the population) and earnings gap (40,200 DKK versus 46,500 DKK in the population) are also similar. The reach–safety earnings gap is the difference between the average earnings of students admitted to their reach program and the average earnings of students admitted to their safety program, and is on average 16,000 DKK. The reach–safety gender gap is the difference between the earnings gender gap for students admitted to their reach program and the earnings gender gap for students admitted to their safety program, and is on average 1,100 DKK with substantial variations. In the next section we decompose the earnings and gender earnings gaps by field. Appendix Tables A1 and A3 present the descriptive statistics separately for the subsamples of students who apply through the Quota 1 and Quota 2 systems of admission, respectively.

2.6 Earnings and Rankings of Fields of Study

Although our analysis exploits the variation in admission at the program level, we summarize the data in broad fields of study. Each program is categorized by Statistics Denmark into educational areas, such as agriculture, arts, mathematics, architecture or basic university programs.[21] We regroup these categories into 10 broad fields of study. The left panel of Table 2 presents descriptive statistics for all students – not just those in the reach-safety sample – who were admitted to a variety of fields of study (all applications; left). This panel summarizes the variation across programs – within and between fields – in the average earnings and gender earnings gaps that we use to identify the causal effects of gender earnings gap. The standard errors show the variation over programs within a broad field. In particular, it is interesting to see that the correlation between earnings and the gender earnings gap is not perfect. For example, law is the highest paid field but has one of the lowest relative gap at 13%. However, the three next highest paying fields (business, engineering, and technology) have some of the highest gaps, both relatively and absolutely. This shows that there is ample variation in earnings and earnings gaps in the data to identify our parameters. Women tend to be more numerous in fields where the gender gap is lower (as is the case in Education, Health and Humanities). One notable exception is business, where the share of share of women in business is almost half despite the fact that the field has the biggest absolute

[21] We exclude vocational programs from the analysis as they are usually not comparable to university programs.

gender earnings gaps.

The right panel of Table 2 presents the relative frequency at which fields are listed as reaches or safeties by women and men in our reach–safety programs pair sample (used sample; right). Although most fields are ranked with similar frequency by men and women, there are a few notable exceptions: men are more likely than women to rank programs in business as their reach and safety choices; women are more likely than men to rank programs in education and health as their reach and safety choices. Note that, in general, women are more likely to rank fields of studies where the gender earnings gap is low. Because our identification strategy focuses on randomizing students across these reach and safety programs, we elicit the returns to education while controlling for selection issues in admissions. In Section 5.2, we further show that our results are robust to including fixed effects for reach and safety pairs.

3 Identification

This study aims to estimate the following econometric model to predict an individual's earnings using educational variables available to researchers:

$$y_i = \delta \overline{y_{p_i}} + \lambda \overline{g_{p_i}} + \beta X_i + \epsilon_i, \tag{2}$$

where y_i denotes the earnings of student i who enrolls in program p_i , $\overline{y_{p_i}}$ denotes the average earnings of individuals who enroll in program p_i, $\overline{g_{p_i}} \equiv \overline{y_{p_i,m}} - \overline{y_{p_i,w}}$ denotes the average gap between the earnings of men and women who enroll in program p_i, X_i denotes individual covariates (including a constant) and ϵ_i is the error term.[22]

We are interested in recovering and interpreting λ, the degree to which the gender earnings

[22] In our baseline specification, earnings are annualized averages over the period 10 to 12 years after initial enrollment. Note that average earnings in a program will be a weighted average of the earnings for men and women who enroll in the program ($\overline{y_{p,m}}$ and $\overline{y_{p,w}}$, respectively), with weights proportional to the share of men and women in that program. Because programs include a relatively large number of individuals enrolled over many years, including or excluding y_i from the calculation of $\overline{y_{p_i}}$ has virtually no effect on the value of $\overline{y_{p_i}}$ or the results that follow from it. For individuals who are not admitted to any program of study, we calculate the average earnings of their enrollment cohort as the sample average of students who were not admitted to any program in the reach–safety programs pair sample used in our main analysis. Therefore, these averages potentially include both the effect of enrolling in a given program and the effect of program peers. To address this concern, in Section 5.4, we test the robustness of our main results to alternative ways of calculating average earnings, first by excluding the enrollee's own cohort year from the calculation, and second by excluding the enrollee's own cohort year and any ulterior cohort year from the calculation.

gap of a program affects a prospective student entering this program, keeping average program earnings $\overline{y_{p_i}}$ constant. λ measures the increase in earnings an individual receives by enrolling in a program that has a \$1 larger gender earnings gap; this coefficient will be estimated separately for men and women. An OLS regression of equation (2) may not reveal an unbiased estimate of λ in the presence of selection problems, that is, if $\text{cov}(\overline{y_{p_i}}, \epsilon) \neq 0$ or $\text{cov}(\overline{g_{p_i}}, \epsilon) \neq 0$. For example, unobserved ability could correlate with program admission.

We instead obtain a causal estimate of λ from equation (2) using the discontinuity built into the Danish university admissions system described in Section 2.3. The natural null hypothesis is $\lambda = 0$; enrolling in a program with a larger gender earnings gap does not causally affect the individual's earnings. In this case, cross-program differences in average enrollee earnings merely reflect selection or the tendency of high- and low-ability students to enter different programs. One alternative hypothesis is that cross-program differences in the gender earnings gap are due to heterogeneous program returns for men and women, with no cross-program differences in selection. Section 3.1 shows that if all programs had equal numbers of men and women, this would correspond to the hypothesis that $\lambda = 0.5$ for men and -0.5 for women. Because the share of men and women varies across programs, a better way to measure selection is to test the equality of λ with the OLS estimate: the OLS and true coefficients will coincide when there are no selection effects.

We estimate a fuzzy regression-discontinuity IV design, in which we use instruments to predict the average earnings and gender earnings gaps of the program the student enrolls in ($\overline{y_{p_i}}$ and $\overline{g_{p_i}}$, respectively). The following variables are included as instruments in Z:

$$\mathbb{P}(p_i = r_i | D_i) \times \overline{y_{r_i}}; \mathbb{P}(p_i = s_i | D_i) \times \overline{y_{s_i}}; \mathbb{P}(p_i = r_i | D_i) \times \overline{g_{r_i}}; \text{and } \mathbb{P}(p_i = s_i | D_i) \times \overline{g_{s_i}}. \tag{3}$$

These instruments can be used in a fuzzy-discontinuity design because they include the probabilities of admission to the reach versus safety program choice. The probability of being admitted to the reach and safety programs therefore respectively increase and decrease discontinuously when comparing scores just below and just above the reach program choice cutoff score. To verify their validity, we investigate bunching and other density tests in Section 3.2. Here, we present the first-stage regressions given by:

$$\overline{y_{p_i}} = \phi X_i + \chi Z_i + \epsilon_i \tag{4a}$$

$$\overline{g_{p_i}} = \phi_g X_i + \chi_g Z_i + \epsilon_{g,i}. \tag{4b}$$

We use the predicted values from the fuzzy discontinuity in these first-stage regressions as a source of variation to obtain unbiased estimates of δ and λ in equation (2). Figure 2 provides a graphical representation of first-stage equations (4a) and (4b), replacing a continuous measure of probability of enrollment in the reach or safety program with a sequence of score bins. This allows us to see graphically the discontinuous change in the average earnings and gender earnings gap of the enrollment program as the admission cutoff is crossed, depending on the earnings or gender earnings gaps of the reach and safety programs. We construct 10 such dummy bins, as cutoff levels are binned to 0.1 centered around 0. Each student will have exactly one dummy bin variable equal to 1, with the rest equal to zero. The reach and safety coefficients are shown in the top panels of Figure 2 for average earnings and the gender earnings gap, respectively. The reach coefficients jump up sharply as D_i goes from negative to zero/positive; once the students' scores are high enough for admission to their reach programs, the average earnings of the reach programs strongly influence the average outcome of the program in which they enroll, as this is typically the reach program. Similarly, the safety coefficients fall sharply as D_i goes from being negative to zero/positive; once the students' scores are high enough for admission to their reach programs, the average earnings of the safety program no longer strongly influence the average outcome of the program in which they enroll, as this is typically no longer the safety program.

Note that these jumps appear discontinuous, are large, and are estimated very precisely. This reflects the fact that the Danish admissions system does in fact operate as described, and thus moving across the grade threshold will frequently land students in their reach program instead of their safety program. Concerns about weak instruments are absent here. Appendix Figure A2 presents the reduced form results for the Quota 1 admission system and Appendix Figure A3 presents the first-stage and reduced form results for the Quota 2 system.

3.1 Signal Extraction Problem

To understand equation (2), consider a more fundamental model to predict the same individual's earnings but with variables that may be unobservable:

$$y_i = a_i + c(p_i) + bX_i + e_i. \tag{5}$$

The individual's earnings y_i can be decomposed into four parts: (i) a_i, the component attributed to the ability of the individual; (ii) $c(p_i)$, the causal component, the return the student receives from enrollment in program p_i, which may differ between men and women for a given program (in which case, $c(p,m) \neq c(p,w)$); (iii) X_i, observed characteristics affecting earnings; and, (iv) e_i, an error term.

Ceteris paribus, prospective students may want to identify programs with higher returns $c(p)$ or where the returns for men and women differ ($c(p,m) - c(p,w) \neq 0$). However, information about $c(p)$ and $c(p,m) - c(p,w)$ is not directly observable. Instead, we only observe the average earnings for students who enrolled in each program ($\overline{y_p}$) and the gender earnings gap for students who enrolled in each program ($\overline{g_p}$). These observables are the sum of the causal return component of interest and the ability or selection component that we may wish to filter out.

If the distributions of $c(p)$ and $\mathbb{E}[a(p)|\overline{y_p}]$ are jointly normal over program p (and for ease of exposition here, uncorrelated with each other and other observables X), the solution to this signal extraction problem is:[23]

$$c(p_i) = \bar{c} + \frac{v_c^2}{v_c^2 + v_a^2}(\overline{y_{p_i}} - \bar{c}) + e_{p_i} \tag{6a}$$

$$c(p_i,m) - c(p_i,w) = \overline{c_m} - \overline{c_w} + \frac{v_{cg}^2}{v_{cg}^2 + v_{ag}^2}(\overline{g_{p_i}} - (\overline{c_m} - \overline{c_w})) + e_{g_{p_i}}. \tag{6b}$$

Plugging equation (6) into equation (5) (when average earnings and the gender earnings gap are uncorrelated and the shares of men and women in each program are equal) yields equation (2),

[23] Let $\bar{a} \equiv \mathbb{E}[a]$, $\bar{c} \equiv \mathbb{E}[c(p_i)]$, $\overline{c_m} \equiv \mathbb{E}[c(p_i,m)]$, $\overline{c_w} \equiv \mathbb{E}[c(p_i,w)]$, $v_c^2 \equiv \text{var}[c(p_i)]$, $v_{cg}^2 \equiv \text{var}[c(p_i,m) - c(p_i,w)]$, $v_a^2 \equiv \text{var}[a]$, $v_{ag}^2 \equiv \text{var}[a_m - a_w]$ and $\text{cov}_{ca} \equiv \text{cov}[a, c(p_i)] = 0$ (as well as all other covariances equal to zero by assumption). While we ignore non-zero correlation terms, these can be accommodated at the expense of expositional simplicity. e_{p_i} represents the program-specific error, the degree to which a program has a higher or lower causal effect than would be predicted by the average earnings of its enrollees.

but with $\delta = \frac{v_c^2}{v_c^2+v_a^2}$ and $\lambda = \pm\frac{1}{2}\frac{v_{cg}^2}{v_{cg}^2+v_{ag}^2}$, with differing signs for men and women. This makes clear that if $v_a^2 = 0$, that is, if there is no variation across programs in the the ability gap between women and men, then $\lambda = \pm 0.5$. In such case there is no selection since it is the variation in differential returns for women and men that cause all of the gender earnings gap. This also shows that if $v_c^2 = 0$, that is, if there is no variation across programs in the differential returns to education for women and men, then $\lambda = 0$. In this case, all variation in the gender earnings gap reflects selection into the programs. Intermediate values are possible when the returns to education reflect partly selection into programs, and partly heterogeneity across women and men. Note that the parameters estimated separately for men and women need not be of equal and opposite sign when estimated using actual data. This could be the case when men and women select from different groups of programs for which different parameters are appropriate, or when programs that tend to have causal gender gaps for women tend to have positive selection for men (or vice versa).

3.2 Manipulation Tests

When prospective students do not know the exact admissions cutoffs, they have a strong incentive to honestly rank their preferences. These preferences should not jump discontinuously around a cutoff that is unknown to applicants. As a result, students should not "bunch" just above the cutoff of their reach program; the number students and the attributes of those students should be similar for students with scores just above and below their reach program's cutoff. McCrary (2008) provides a manipulation test to look for this bunching the context of RDD. The idea is that absent manipulation, the density of units around the admission cutoff should be continuous. In our case, although the assignment variable is discrete (grades and cutoffs are calculated in 0.1 increments), Lee and Lemieux (2010) suggest that local linear or polynomial regressions can be used to test jumps around the cutoff, as would be the case for a continuous assignment variable. The main differences in our case are that the data are naturally binned in 0.1 increments, and that the choice of bandwidth around the cutoff is therefore more limited. We investigate manipulation of the assignment variable by looking at bunching around the cutoff, year-to-year dynamics of the cutoff, and covariate balance around the cutoff.

3.2.1 Bunching

The density of the assignment variable D_i is presented in Figure 3 (a). We present the number of applicants in each score bin around the cutoff, and segment the total number of applications into women and men. Initial visual inspection provides evidence of bunching around the cutoff for neither men nor women. Nonetheless, a manipulation test proposed by Cattaneo et al. (2019) can be used to test for the continuity of the density function of D around the normalized cutoff point 0. Formally, we are interested in testing

$$H_0 : \lim_{d\uparrow 0} f(d) = \lim_{d\downarrow 0} f(d), \tag{7}$$

against the alternative that the two limits are not equal. Following Cattaneo et al. (2019), a test statistic can be constructed using a local-polynomial density estimator based on the c.d.f. of the observed sample which takes the form:

$$T_p(h) = \frac{\hat{f}_{+,p}(h) - \hat{f}_{-,p}(h)}{\hat{V}_p(h)} \sim N(0,1) \qquad \hat{V}_p^2(h) = \hat{\mathbb{V}} = \left\{\hat{f}_{+,p}(h) - \hat{f}_{-,p}(h)\right\}. \tag{8}$$

The test statistic depends on the local-polynomial density estimators $\hat{f}_{+,p}(h)$ and $\hat{f}_{-,p}(h)$, and on the standard error estimator $\hat{V}_p(h)$. The parameter h denotes the bandwidth used around the cutoff point, while p denotes the choice of polynomial order. Additional inputs to compute the test statistic include the choice of standard errors (either conventional or jackknife), and the restrictions imposed on the model. The unrestricted version of the test allows both estimators $\hat{f}_{+,p}(h)$ and $\hat{f}_{-,p}(h)$ to be unrelated. The restricted version of the test is more powerful but assumes that the c.d.f. and higher-order derivatives of the running variable are equal for treated and untreated groups at the cutoff even when $f_- \neq f_+$.

Because of the nature of the data, we choose to present the results of the test using local-linear regressions for varying bandwidths around the cutoffs, under both the restricted and unrestricted tests and using both conventional and jackknife standard errors. Table 3 presents the p-values for the hypothesis test in equation (7) under these different cases. Under all of these cases, we reject manipulation of the assignment variable at the 5% statistical significance level. This provides evidence that the admission system works and that individuals are quasi-randomly assigned to different university programs based on their score.

3.2.2 Dynamics of Cutoffs

Variation in admission cutoffs is necessary for students to truthfully reveal their preferences. Figure 3 (b) shows a histogram of the difference between the admission cutoff for the reach program in the student's year of application and the admission cutoff for the same program in the previous year.[24] To further investigate how cutoffs move from year to year, Figure 3 (c) presents the same data but conditional on the students' distance from the admission cutoff of the reach program in the year prior to their application. It shows that students who were between 0.2 points below and 0.1 point above the cutoff for their reach program based on the cutoff in the previous year face similar levels of uncertainty in relation to cutoff changes.

This provides evidence that prospective students face substantial annual changes in cutoffs, making it hard to game the system by guessing the admission cutoff for their most preferred program based on past cutoffs. Nonetheless, to consider the possibility that students base their application decision on the admission cutoff from the previous year, we construct the probability of being admitted into the reach and safety programs based on a student's current grade and the previous admission cutoffs. Our regressions control for these probabilities, interacted with the average outcomes for the reach and safety programs. This allows us to identify the variation in unanticipated changes in score cutoffs in the main analysis.

3.2.3 Covariate Balance

Finally, if manipulation around the cutoff is impossible by prospective students, we should not find discontinuous changes in observable characteristics of the applicants and of programs around the cutoff. We verify this in Figure 4, which presents the average of covariates used in the analysis. We start by looking at variables that affect admission into the reach program. Panel (a) shows the high school score used to rank students when considering admission to programs. The score trends up linearly around the cutoff without any jump at the cutoff. Panel (b) shows the total

[24] We restrict the histogram to changes in program admission cutoffs that are between –1 and 1 because they account for most of the data; 1,262 observations are for programs that have an increase in the admission cutoff of greater than 1, and 29 observations are for programs that have a decrease in the admission cutoff of less than –1.

number of programs ranked which trends down, while Panel (c) shows the proportion of admissions that have the same field for the reach and the safety and there is also no jump apparent. Finally Panel (d) shows parental wealth as measured the year before university admission and again no jump is apparent. Finally, Panels (e) and (f) show the average earnings and earnings gap for reach and safety programs. None of these characteristics of students, their application or their reach-safety programs suggest manipulation around the cutoff.

4 Results

4.1 Main Results

Table 4 presents our main results. The top panel shows results from the fuzzy-discontinuity IV design, while the lower panel shows results from an OLS estimation on the same sample, controlling for a set of observables.[25] The OLS estimations provide an intuitive benchmark for our results because they represent the best forecasts of prospective students or researchers that do not observe the admission discontinuity. However, because they are cross-sectional averages, they represent correlations and cannot be interpreted in a causal way; they potentially include a selection effect. For this reason, our causal estimates from the IV estimation show what part of the correlation can be attributed to the causal effects of the returns to education.

In the first column we estimate a bare bone model with only program earnings, i.e. setting $\overline{g_{p_i}}$ to 0 in equation (2). We instrument the average earnings of the program the student enrolls in ($\overline{y_{p_i}}$) with equation (4a) using reach and safety program earnings, the two instruments from equations (3a) and (3b). Our fuzzy-discontinuity IV estimate $\hat{\delta} = 0.26$ (C.I. 0.055 to 0.47) implies that entering a program whose enrollees earn \$1 more than other programs leads to a \$0.26 increase in earnings on average. A coefficient of $\delta = 0$ would suggest that cross-program variation in average earnings entirely reflects selection, whereas a coefficient of $\delta = 1$ would reflect a perfect causal effect of the program studied. Because this result is far less than 1, and even less than 0.5, it implies that most of the cross-program differences in average earnings reflect selection, even though we find statistically significant causal effects. The OLS benchmark shows a much larger effect of 0.80, providing evidence that the returns to education one can observe without random variation in

[25] The control variables are parental wealth, income and education, high school score, age and year fixed effects.

admissions largely represent selection. Our results therefore suggest that enrolling in a program with higher earnings will lead to increased earnings, but that most cross-program variation in average earnings reflects selection and will lead to differences in program returns.

In the second and third columns of Table 4, we estimate equation (2) separately for women and men, instrumenting average earnings and the gender earnings gap of the student's admission program with the equations in (4) using reach and safety earnings and gender earnings gaps, the four instruments from the equations in (3).

The first row of Columns 2 and 3 of Table 4 show that the causal effect of program earnings is stronger for women than men. The IV estimates of $\hat{\delta} = 0.48$ (C.I. 0.23 to 0.74) for women and $\hat{\delta} = 0.06$ (C.I. –0.31 to 0.43) for men show that average program earnings impact earnings for women but not for men. The divergence of OLS results for men (0.75, C.I. 0.65 to 0.85) from their IV estimates suggest that cross-program differences in earnings for men reflect selection of high-ability men into programs that consequently show high earnings. Conversely, the OLS benchmark of 0.79 (C.I. 0.73 to 0.85) for women is much closer to the IV estimate, suggesting that cross-program differences in earnings for women reflect cross-program differences in returns to education.

Columns 2 and 3 of Table 4 next show the main results of the paper: women obtain lower returns when they enroll in programs with higher gender earnings gaps, holding average program earnings constant. Our fuzzy-discontinuity IV estimate $\hat{\lambda} = -0.45$ (C.I. –0.81 to –0.096) implies that entering a program with a $1 larger gender earnings gap leads to a $0.45 reduction in earnings for women on average. Interestingly, this effect is stronger than the analog benchmark OLS result. This could be the case if women were less likely to self select into programs with large gender earnings gaps, as is suggested by the results of Table 2. This implies that the causal effect of gender earnings gap is actually larger than what one would estimate using only observational data.

Because we reject the hypothesis that gender earnings gaps do not affect earnings for women, we can also reject the idea that cross-program differences in gender earnings gaps merely reflect differential selection of high- and low-ability men and women into different programs. The analogous coefficient for men in the third column $\hat{\lambda} = -0.09$ (C.I. –0.58 to 0.41) is statistically indistinguishable from zero and can rule out large positive values.[26] For men the OLS benchmark results actually suggest a positive premium to studying programs in which the gender earnings gap is larger, but this effect does not translate to our IV results that control for selection. In this case, although we observe men earning more in fields with larger gender earnings gap, it is not the case once we account for selection.

In the fourth and fifth columns of Table 4, we estimate equation (2) separately for women and men using gender-specific program earnings. We instrument average gender-specific earnings with gender-specific reach and safety earnings. These present the same results in a different way; women earn significantly more when they enroll in programs with higher-earning women (holding male earnings fixed), with a point estimate of 0.76 (C.I. 0.25 to 1.27). We can reject the hypothesis that the earnings of men and women in a program affect a woman's earnings equally. By contrast, the fifth column shows that male and female earnings in a program are equally important, although noisy, in relation to a man's earnings. The benchmark OLS results show that both women and men earnings mostly load on their gender-specific program earnings.[27]

4.2 Alternative Channels

This paper has documented that programs with larger gender earnings gaps reduce earnings for women. In this section, we investigate potential mechanisms through which this result may operate.

Programs could lead to women earning less than men through direct channels, such as explicit discrimination, or through indirect channels. An important body of literature documents an earnings penalty for mothers but not fathers. For example, Kleven et al. (2018) finds that part of the gender wage gap can be explained by women having children. In this case, one possible explanation of our findings could be that some fields have a causal impact on women's decisions

[26]Note that the parameters estimated separately for men and women need not be equal and of opposite sign – as suggested in the identification section – when estimated using actual data. This could be because men in the reach–safety programs pair sample may apply to a different set of programs than women, for which the causal effect of the gender earnings gap is small or absent, or because programs that tend to have large causal gender gaps for women tend to have positive selection in relation to male enrollees, thereby offsetting the causal effect of the gender gap for men. Nevertheless, in both of these cases, women do worse when they enter programs with large gender earnings gaps.

[27]Appendix Table A5 shows the first-stage estimations of equation (3) for the results presented in Table 4. The coefficients on these instruments have high predictive power: they are close to 1 for average reach and safety program incomes, and the R-squared values for the regressions are high and vary between 79% and 94% as suggested by their graphical representation in Figure 2.

to have children, and that women in these fields earn less as a result of the earnings penalty documented for motherhood.[28] In this section, we analyze whether cross-field/program differences in the gender wage gap (not the overall gap) can be explained by marriage, number of children, graduation rates and the labor market participation of women.

To explore these possible channels, Table 5 shows the baseline results and includes additional controls. The first column reproduces our baseline result for women. In the second column, we include as controls a dummy variable indicating whether the woman is married 10 to 12 years after enrollment and the number of children she has during the same period. Our results are robust to the inclusion of these controls, As a result, we rule out the hypothesis that programs with large gender earnings gaps have them because they lead women in those fields to get married or have more children, which in turn leads to lower earnings.[29]

In the fourth column of Table 5, we include a dummy variable indicating whether the woman has graduated from her reach program 10 to 12 years after enrollment. Including a a graduation control has almost no impact on our point estimates, providing no evidence that a graduation channel explains our gender earnings gap findings. By including graduation controls, we rule out a channel through which programs with large gender earnings gaps have them because they lead women in those fields to graduate at lower rates, which in turn would lead to lower earnings. Including a dummy for the presence of any labor income modestly (but not significantly) reduces the estimated impact of a program's gender earnings gap on the earnings of its female enrollees. This might provide some suggestive evidence that programs with large gender earnings gaps lead to women earning less on the extensive margin (not working), as well as on the intensive margin (earning less while working).[30]

[28]Gallen et al. (2017) also finds that the gender gap in Denmark has decreased over the last 30 years, with a further reduction linked to higher female labor market participation. However, the remaining gender gap is still related to the child penalty.

[29]Furthermore, in unreported results, we estimate analogous IV equations to predict marriage and number of children and find no causal effect of the program studied on either of these outcomes (with p-values ranging from 0.29 to 0.81 for our coefficients of interest).

[30]In unreported results, we estimate analogous IV equations to predict graduation and the presence of any labor income, and find very modest effects of average program earnings on graduation rates; an increase of 0.1 percentage point (C.I. 0.018pp to 0.19pp) per increase of 1,000 DKK and of a gender earnings gap in the presence of any labor income and a reduction of 0.04 pp (C.I. –0.10pp to 0.001pp) per increase of 1,000 DKK.

4.3 Gender Earnings Gap Dynamics

There is growing evidence that the difference in earnings between women and men at the beginning of their careers is small, but grows to be considerable a decade after graduation (Bertrand et al., 2010, for example). We verify this hypothesis by re-estimating our baseline results while varying the time horizon over which earnings are averaged. Table 6 shows the results using individual and program earnings measured 7 to 9 years, 10 to 12 years (our baseline), and years 13 to 15 after enrollment. Panel A shows the main IV results, and Panel B shows the OLS results for comparison.

The results show that the causal effect that we documented in Table 4 is not present seven to nine years after enrollment: the within-program gender earnings gap estimate is both much smaller and statistically insignificant. However, the causal estimate we documented in the main specification for years 10 to 12 (i.e. coefficient –0.45, s.e. 0.18), is also statistically significant in years 13 to 15, and even larger (i.e. coefficient –0.80, s.e. 0.23). This provides evidence that cross-program differences in the gender earnings gap tend to increase with time spent in the labor market.[31] We therefore contribute to the evidence on gender earnings gaps growing with the employment life-cycle.

4.4 University, Field, and Program Effects

We investigate the roles of gender earnings gaps at the level of the academic institution (universities), broad program of study (e.g., business versus humanities), and specific program of study in affecting subsequent earnings.

We start by decomposing the gender earnings gap of a program into the university-wide component (the average for all programs at that university) and the program-specific component (the residual contribution for that program compared to others within the same university). We then reproduce our baseline results but decompose the average program earnings $\overline{y_p}$ into these program- and university-specific averages. This allows us to separate out the effect of the program and the university of study. If universities, rather than programs, are driving the results for the gender earnings gap, we should find that university-wide gender earnings gaps have the

[31]Table A6 repeats the estimations at different time horizons after enrollment, as shown in Table 6. We find that the result that the gender gap increases over time remains even after controlling for children and other important controls.

largest causal effect.

The two first columns of Table 7 decompose gender earnings gaps into the earnings impact of entering a university with a large gender earnings gap versus the impact of entering a program with a relatively large gender earnings gap compared with others in that university.[32] The results are estimated separately for men and women. The second column shows that our gender earnings gap results are present at both program and university levels. Entering a program that has a relatively large gender earnings gap compared with other programs in the university will significantly reduce earnings for women (coefficient –0.42, s.e. 0.18), and entering a university that has a large gender earnings gap will reduce earnings for women even more (coefficient –0.76, s.e. 0.30). Almost all differences across universities in terms of their gender earnings gaps reflect causal effects, with no evidence of selection. As with the overall effects, the gender earnings gap of a program has no effect on earnings for men.

The third column separates the impact of entering a higher-earning university and entering a higher-earning program within that university.[33] While both effects are present, the impact on earnings of entering a higher-earning university (coefficient 0.57, s.e. 0.16) is significantly larger than the impact on earnings of entering a higher-earning program within a university (coefficient 0.17, s.e. 0.11). Roughly half of the differences between universities reflect the causal effects of entering those universities; therefore, the remaining half must reflect the students who select those universities. Finally, the last column re-estimates the baseline regression on the subsample of students whose reach and safety programs are within the same university; in this sample, all variation *across* programs is also *within* universities (e.g. moving from physics to chemistry or economics within the University of Copenhagen). The results in this case are broadly similar, but with a smaller point estimate (coefficient 0.13, s.e. 0,16).

Next, we decompose the gender earnings gap of a program into the field component (the average for all programs in that broad field of study, such as business or humanities) and the program-

[32]Here, we use six instruments: the probability of entering the reach and safety programs, interacted with the average earnings of those programs, the gender earnings gap of the university in which those programs are taught, and the gap between the program's gender earnings gap and the university's gender earnings gap. These instruments are used to predict the average program earnings, the gender earnings gap of the university the student attends, and the gender earnings gap of the student's program relative to that of the university.

[33]Here, we use four instruments: the probability of entering the reach and safety programs, interacted with the average earnings of the university at which those programs are taught, and the gap between the program's average earnings and the university's average earnings. These instruments are used in the first stage to predict the average university earnings and the gap between the program's earnings and the university's earnings.

specific component (the residual contribution for that program compared to others within the broad field). Table 8 reproduces our baseline results but decomposes the gender earnings gaps $\overline{g_p}$ and average program earnings $\overline{y_p}$ into these program- and field-specific averages. This allows us to disentangle the effect of the exact program studied (e.g. economics or finance) versus the broad field of study (e.g. business, health or science).

The two first columns of Table 8 decompose gender earnings gaps into the earnings impact of entering a broad field with a large gender earnings gap versus the impact of entering a specific program with a relatively large gender earnings gap compared with other programs in that field. Both the gender earnings gap in the broad field and that in the specific program matter in relation to women's earnings. Women earn less when they enter broad fields with larger gender earnings gaps (coefficient –0.34, s.e. 0.22) and specific programs with larger gender earnings gaps than that in their broad field (coefficient –0.55, s.e. 0.24). As before, the gender earnings gap – regardless of decomposition – does not affect earnings for men.

The third column decomposes program-specific average earnings into the earnings of the broad field versus the difference between the program's earnings and the broad field's earnings; both types of variation are broadly similar in magnitude and statistically different from zero. This suggests that not only the broad field of study but also the exact program a student chooses are important for earnings. This is confirmed in the last column, which repeats the baseline regression on the subsample of students whose reach and safety programs are within the same university. We find that moving to a higher-earning program within a broad field leads to a significant increase in earnings that is similar to our baseline results (including moves between and within broad fields). These results are quantitatively smaller and statistically indistinguishable from zero, suggesting that cross-program differences are important for earnings.

4.5 Program Gender Composition

Beyond differences in the gender earnings gap, programs differ in the proportion of men and women who enroll. The Pew Research Center conducted a survey and found that "women employed in majority-male workplaces are more likely to say their gender has made it harder for them to get ahead at work, they are less likely to say women are treated fairly in personnel matters, and they report experiencing gender discrimination at significantly higher rates" (Parker, 2018).

Table 9 uses the same identification strategy as before, but exploits the switch of students from programs with fewer men to those with more men as they cross the score threshold. We add two additional instruments (the probability of admission to the reach and safety programs multiplied by the share of men in the reach and safety programs, respectively) and one additional outcome predicted by those instruments in the first stage (the share of men in the program in which the student enrolls). This allows us to estimate the causal effect on earnings of entering an otherwise identical program with a higher share of men. The first column repeats our baseline results in relation to earnings. The second and third columns restrict the sample used to women and men, respectively, and include the share of men in the program of admission.

We find that the share of men in a program has no significant impact on female earnings. Moving a woman from a program with all women to a program with no women would reduce her earnings by 22,000 DKK, or less than 3,400 USD based on our point estimates, but this difference is not close to statistically different from zero (coefficient –21.87, s.e. 32.46). For men, we find that entering a program with a larger share of men leads to a fairly large and statistically significant reduction in earnings (118,000 DKK or approximately 15,000 USD from moving from an all-female program to an all-male one). One possible explanation is that men in programs with a larger share of women receive additional attention and consequently gain higher returns in their field of study.

4.6 Discussion

The results presented in this section contribute to an important discussion on the effect of university programs on gender earnings gaps (Heckman, 1979, Gronau, 1974). Causal effects reflect the challenges and experiences of women in particular fields, while selection reflects differences in the ability and attributes of men and women who enter these fields.[34] While they do not consider the gender earnings gap explicitly, Hastings et al. (2013) compare the causal effect of entering a preferred broad field over all other less preferred options separately for men and women and

[34]In the overall labor market, the evidence on the effects of selection in wage gap estimations is mixed. For example, Blau and Beller (1988) found that the wage gains of white women relative to men were larger after correcting for selection, while Blau and Kahn (2006) and Mulligan and Rubinstein (2008) showed that the wage gap was larger after controlling for selection and that in some cases the gap had not reduced at all.

find gender differences in earnings effects from entering health programs in Chile, but no significant gender differences for other broad fields.[35] Buckles (Forthcoming) reviews studies seeking to address the issue of causality using interventions to reduce the gender wage gap. The common approach in this line of study is to design a randomized control trial to assess which kinds of interventions motivate more females to study in a field in which they are under-represented (Avilova and Goldin, 2018). Moreover, Bennedsen et al. (2019) exploit a 2006 legislation change in Denmark that requires firms to provide wage statistics by gender, finding that it has had an effect in reducing the gender wage gap with no significant overall effects on firm profitability.[36]

Our main results show that a large share of the gender earnings gap across programs arises from differential returns to education for men and women rather than self-selection of men and women of different abilities into different fields. This result is important for policy makers as it suggest earnings discrimination against women. Although we document the presence of gender earnings gap along many dimensions, we are still agnostic about the exact features of programs that cause these gaps. The fact that gender earnings gap are present both at the university and field levels suggests that education institutions have a role to play in preventing discrimination against their students, whether while they are in school or once they reach the job market. Because we find that the gender earnings gap grows over time, more research is needed to understand why fields differ in the career progression for women.

5 Robustness of Results and Alternative Specifications

In this section we investigate numerous robustness tests and alternatives specifications for the main results we document.

5.1 Treatment Identifiers

The baseline specification presented in Table 4 is correct to the degree that initial enrollment in a program is best understood as the treatment. However, students sometimes switch programs,

[35] They interpret this result as consistent with findings in the literature whereby non-pecuniary differences and characteristics across fields are more important in driving choices than pecuniary ones. See, for instance, Goldin and Katz (2008), Bertrand (2011) and Bertrand et al. (2010). For example, women may have preferences related to the proportion of women in the field, or to work schedule and career trajectory that do not overlap with reproductive age.

[36] The main channel was by means of slowing wage growth for male employees.

and students who initially enroll in their safety program are more likely than others to switch. Students' earnings may be affected not only by what they enroll in initially but by what they eventually go on to study. Appendix Figure A5 shows the share of those who are enrolled in their reach, safety, or other program 1, 2,..., 6 years after initial enrollment.[37] Note that the shares of students who enroll in reach and safety programs jump discontinuously across the cutoff threshold. While the size of this jump is slightly higher when looking at the program of initial enrollment, the jump is significant even when looking at the program of enrollment four or six years after initial enrollment. Appendix Figure A6 shows that the attenuation in the jump at the cutoff measured by the program enrolled in many years from initial enrollment reflects an above-average switch rate for students initially enrolled in their safety program choice. Appendix Figure A7 shows that roughly 10% of those enrolled in their reach program eventually switch to their safety program; most of those who switch do so to other programs.

Alternatively, to the degree that the true treatment is time spent studying in a program or graduating from a program, the IV results based on initial enrollment will yield results that are too small (because the first-stage results of movement across the cutoff on the program studies will be too big). Panel A of Table 10 presents the results shown in Table 4 in a different way. First, we show the first and second stages, and then the IV estimates from the baseline specifications in Table 4, along with the OLS estimates. Panels B and C of Table 10 repeat the baseline results of Panel A, but replace the initial-enrollment-based measure of $\overline{y_{p_i}}$ with an average of the earnings over the different programs they studied during the four first years after initial enrollment (Panel B) and the earnings of all those who enrolled in the program the person studied in year four (Panel C). These results are consistent with our understanding of the treatment. Many students who enter a given program eventually switch. The causal effect of entering a program will naturally be smaller than the causal effect of completing it, because much is learned along the way. Our instrument increases the probability of entering the reach program more than it increases the probability of continuing in the reach program, if only because students routinely switch programs. These results echo the graphical representation in Appendix Figure A8. We note that the coefficients are roughly 50% higher, reflecting the fact that those who initially enroll in their safety program switch out of it at

[37]This figure ignores information on drop-outs and considers a student as remaining enrolled in a program until he or she switches to a new program.

somewhat higher rates than other enrollees. We also note that the effect of the gender earnings gap is present under both definitions of average earnings in Panels B and C.

Appendix Figure A8 replaces this measure of $\overline{y_{p_i}}$ with an average of the earnings over the different programs studied during the first four years after enrollment (top panel) and the earnings of all those who enrolled in the program the person studied in year four (bottom panel). While the results are broadly similar, the jump in the coefficients are smaller, suggesting that moving over the cutoff threshold has less impact on the earnings from the programs in which students eventually study than those from the ones in which they initially study. This is unsurprising because students may switch programs after initial assignment as their interests change or they are able to enter a program they had previously been unable to enter.

5.2 Instruments

Table 11 shows the results of robustness tests of the results shown in Table 4. Panel A shows our baseline specification, which controls for distance from cutoff fixed effects (bin FE), average earnings of the reach program, average earnings of the safety program, calendar year fixed effects, and the probability of being accepted given last year's admission cutoff interacted with the average earnings of the reach and safety programs. Panel B adds reach and safety programs fixed effects. Panel C omits the probability of being accepted given last year's admission cutoff interacted with the average earnings of the reach and safety programs. This specification does not control for past cutoffs that are in the information set of prospective students at the time of application, and therefore is identified only on the difference between a student's score and the cutoff for the reach and safety programs. Panel D adds the reach and safety program cutoffs and the student's score. This allows us to verify whether the competitiveness of programs and the ability of students impacts our estimates. Panel E adds demographics (parents' wealth, income and education, and student's age and gender). This allows us to measure the effect while keeping these demographics constant. The results are similar across specifications.

5.3 Bandwidths around admission cutoffs

Figure 5 shows the robustness of our baseline results to different bandwidths of estimation around the admission cutoff. We retain a symmetric number of bins on each side of the cutoff and tighten the bandwidth from +/– 0.5 bins to +/– 0.2 bins around the cutoff. We present our main coefficients for earnings and the gender earnings gap in Panels (a) and (b), respectively. Although the sample size decreases as the bandwidth is tightened, and consequently the statistical power is reduced, the results are consistent across all bandwidths used.

Table A9 presents our baseline results excluding any prospective students who had a score exactly equal to the admission cutoff of their reach program. Because in such cases these students are randomized into their reach program, we check to make sure that our main results are not affected by their inclusion. The results are quantitatively unchanged.

5.4 Average Program Incomes

The baseline specification presented in Table 4 uses average program incomes calculated over the entire sample of observations. Such averages maximize the use of the available data by using all cohort years, thereby producing more precise averages. However, the results estimated using these averages could include cohort and peer effects. In this section, we present two alternative ways to calculate average program incomes to address this issue. We therefore recalculate the average program incomes for students' reach, safety and enrollment programs by excluding their own application cohort year. This removes the cohort effect from our results.

5.5 Admission System

Our main results presented so far use the subsample of applicants that did not choose the standby option. As discussed in Section 2.1, the cutoff for the standby list is usually 0.1 to 0.2 grade points below the actual cutoff point. If students choose to apply for standby places, they will be offered a standby place instead of their safety program if they are above the cutoff point for the standby list. In case they are not offered a place within the first month of the program starting, they are guaranteed a place in the program the following year. Because the standby option sometimes delays the start of the program, we do not include these applications in our main analysis. Appendix

Figure A1 shows the admission discontinuity for each system.

In this section, we verify that our baseline results are robust to including the standby applications. Table A7 show the results of our analysis using the entire sample of applicants. We enrich our econometric specification to include admission-system fixed effects (i.e. fixed effects for each of the Quota 1, Quota 2, Quota 1 on standby, and Quota 2 on standby admission systems) and their interactions with the distance from the cutoff, as well as the income and gender earnings gaps for the student's reach and the safety programs. We also present the robustness of these results to the bandwidth around the admission cutoff in Panel (a) of Figure A4. In Table A8 we also present the same estimations for the subsample of students that applied through Quota 1 only, whether on standby or not. We also present the robustness of these results to the bandwidth around the admission cutoff in Panel (b) of Figure A4. The results are qualitatively the same then the ones presented in Section 4.

6 Conclusion

We exploit a discontinuity built into the Danish university admissions system that quasi-randomly assigns similar students to different university programs. Prospective university students receive a composite score summarizing their performance in high school. They then rank programs – university–major pairs, such as medicine at the University of Copenhagen – as their first, second, third, and so on, choices. Each program has a score cutoff, which changes from year to year to clear the market. Students are admitted to their most preferred program for which their score exceeds the cutoff. We examine students with reach–safety program pairs wherein their score is close to the cutoff for their more preferred "reach" program but well above the cutoff for their next-most-preferred "safety" program. For these students, a small improvement in their score from just below to just above the cutoff will lead to a large, discontinuous increase (decrease) in their probability of enrolling in their preferred reach (less preferred safety) program.

We find that enrolling in programs with higher-earning enrollees leads to students earning more, but that most cross-program differences in earnings reflect selection. Entering a program whose enrollees earn \$1 more will on average lead to an enrollee earning \$0.26 more (C.I. \$0.055 to \$0.47). This effect is somewhat larger when we redefine the treatment as studying the subject four years after initial enrollment instead of at initial enrollment. The results are robust to a variety of bandwidths and controls. We document substantial and important heterogeneity by gender. In particular, the cross-program differences in the gender earnings gap are causal; enrolling in a program whose enrollees experience a $1 larger gender earnings gap, controlling for earnings, leads to women earning $0.45 less on average (C.I. –0.81 to –0.096). This effect is absent when women initially enter the labor market, but grows over time to $0.80 (C.I. –1.24 to –0.36).

Although documenting cross-program differences in the gender earnings gap is important, more research is required to understand the potential channels through which it operates. For example, we find that it is not the result of cross-program differences in marital status and children, but our analysis cannot explain whether it is the result of challenges faced by women in specific university programs or those faced by women in the jobs that those programs eventually feed into. The fact that we find evidence of effects of gender earnings gap causation at both the field and university levels suggests that they could be related to both specific fields and specific universities. However, more research is required to understand whether the challenges faced by women are determined by the education women attain, by the labor market conditions in the various fields in which women find work, or a combination of both. Identifying the importance of each of these channels is useful for policy-makers designing interventions or policies aimed at reducing the gender earnings gap.

References

ALTONJI, J. G., P. ARCIDIACONO, AND A. MAUREL (2016): "The Analysis of Field Choice in College and Graduate School: Determinants and Wage Effects," *Handbook of the Economics of Education*, 5, 305–396.

ALTONJI, J. G., E. BLOM, AND C. MEGHIR (2012): "Heterogeneity in Human Capital Investments: High School Curriculum, College Major, and Careers," *Annual Review of Economics*, 4, 185–223.

ARCIDIACONO, P. (2004): "Ability Sorting and the Returns to College Major," *Journal of Econometrics*, 121, 343–375.

AVILOVA, T. AND C. GOLDIN (2018): "What Can UWE Do for Economics?" *National Bureau of Economic Research Working Paper 24189*.

BENNEDSEN, M., E. SIMINTZI, M. TSOUTSOURA, AND D. WOLFENZON (2019): "Do firms respond to gender pay gap transparency?" *NBER Working Paper No. 25435*.

BERTRAND, M. (2011): "New Perspectives on Gender," *Handbook of Labor Economics*, 4, 1543–1590.

BERTRAND, M., C. GOLDIN, AND L. F. KATZ (2010): "Dynamics of the Gender Gap for Young Professionals in the Financial and Corporate Sectors," *American Economic Journal: Applied Economics*, 2, 228–255.

BISHU, S. G. AND M. G. ALKADRY (2017): "A Systematic Review of the Gender Pay Gap and Factors That Predict It," *Administration & Society*.

BLAU, F. D. AND A. H. BELLER (1988): "Trends in Earnings Differentials By Gender, 1971–1981," *ILR Review, Volume 2*, 1051–1078.

BLAU, F. D. AND L. M. KAHN (2006): "The U.S. Gender Pay Gap in the 1990s: Slowing Convergence," *Industrial and Labor Relations Review*, 60, 45–66.

——— (2017): "The Gender Wage Gap: Extent, Trends, and Explanations," *Journal of Economic Literature*, 55, 789–865.

BORING, A., K. OTTOBONI, AND P. B. STARK (2016): "Student Evaluations of Teaching (Mostly) do not Measure Teaching Effectiveness," *ScienceOpen Research*.

BUCKLES, K. (Forthcoming): "Fixing the Leaky Pipeline: Strategies for Making Economics Work for Women at Every Stage," *Journal of Economic Perspectives*.

CAIN, A. AND A. KIERSZ (2018): "The 19 Jobs in the US with the Biggest Gender Pay Gaps," [Online; posted 10-April-2018].

CALONICO, S., M. D. CATTANEO, AND M. H. FARRELL (2018): "On the effect of bias estimation on coverage accuracy in nonparametric inference," *Journal of the American Statistical Association*, 113, 767–779.

CARD, D. (1999): "The Causal Effect of Education on Earnings," *Handbook of Labor Economics*, 3, 1801–1863.

CATTANEO, M. D., M. JANSSON, AND X. MA (2019): "Simple local polynomial density estimators," *Journal of the American Statistical Association*, 1–7.

CECI, S., D. GINTHER, S. KAHN, AND W. WILLIAMS (2014): "Women in Academic Science: A Changing Landscape," *Psychological Science in the Public Interest*, 15, 75–141.

DALE, S. B. AND A. B. KRUEGER (2002): "Estimating the Payoff to Attending a More Selective College: An Application of Selection on Observables and Unobservables," *Quarterly Journal of Economics*, 117, 1491–1527.

DALY, M. AND D. LE MAIRE (2019): "University admission and preferred field of study," *Working Paper*.

D.ANGRIST, J. AND G. W. IMBENS (1995): "Two-Stage Least Squares Estimation of Average Causal Effects in Models with Variable Treatment Intensity," *Journal of the American Statistical Association*, 90, 431–442.

EISENHAUER, P., J. J. HECKMAN, AND S. MOSSO (2015): "Estimation of Dynamic Discrete Choice Models by Maximum Likelihood and the Simulated Method of Moments," *International Economic Review*, 56, 331–357.

FOLEY, K. AND F. GROES (2018): "Admissions Constraints and the Decision to Delay University," *Working Paper*.

GALLEN, Y., R. V. LESNER, AND R. VEJLIN (2017): "The Labor Market Gender Gap in Denmark: Sorting Out the Past 30 Years," *IZA DP No. 10789*.

GOLDIN, C. AND L. F. KATZ (2008): "Transitions: Career and Family Life Cycles of the Educational Elite," *American Economic Review*, 98, 363–369.

GOLDIN, C., S. KERR, C. OLIVETTI, AND E. BARTH (2017): "The expanding gender earnings gap: Evidence from LEHD-2000 Census," *American Economic Review: Papers and Proceedings*, 110–114.

GRONAU, R. (1974): "Wage comparisons—A selectivity bias," *Journal of Political Economy, Volume 82*, 1119–1143.

HANUSHEK, E. A. AND S. G. RIVKIN (2006): "Teacher Quality," *Handbook of the Economics of Education, Volume 2*, 1051–1078.

HASTINGS, J. S., C. A. NEILSON, AND S. D. ZIMMERMAN (2013): "Are Some Degrees Worth More Than Others? Evidence From College Admission Cutoffs in Chile," *NBER Working Paper No. 19241*.

HAUSMANN, R., L. TILL, V. RATCHEVA, R. SAMANS, L. TYSON, S. ZAHIDI, AND W. ECONOMIC FORUM (2017): "The Global Gender Gap Report," *http://www3.weforum.org/docs/WEF_GGGR_2017.pdf*.

HECKMAN, J. J. (1979): "Sample Selection Bias as a Specification Error," *Econometrica*, 47, 153–161.

HECKMAN, J. J., J. E. HUMPHRIES, AND G. VERAMENDI (2018): "Returns to Education: The Causal Effects of Education on Earnings, Health, and Smoking," *Journal of Political Economy*, 126, S197–S246.

HEINESEN, E. (2018): "Admission to Higher Education Programmes and Student Educational Outcomes and Earnings – Evidence from Denmark," *Economics of Education Review*, 63, 1–19.

HEINESEN, E. AND C. HVID (2019): "Returns to Field of Study in the Medium Term – Instrumental Variables Estimates Based on Admission Thresholds," *Working Paper*.

HENGEL, E. (2017): "Publishing While Female. Are Women Held to Higher Standards? Evidence from Peer Review," *Cambridge Working Papers in Economics 1753*.

HOXBY, C. M. (2009): "The Changing Delectivity of American Colleges," *Journal of Economic Perspectives*, 23, 95–118.

HUMLUM, M. K., J. H. KRISTOFFERSEN, AND R. VEJLIN (2017): "College admissions decisions, educational outcomes, and family formation," *Labour Economics*, 48, 215–230.

KEANE, M. P. AND K. I. WOLPIN (1997): "The Career Decisions of Young Men," *Journal of Political Economy*, 105, 473–522.

KIRKEBOEN, L. J., E. LEUVEN, AND M. MOGSTAD (2016): "Field of Study, Earnings, and Self-selection," *The Quarterly Journal of Economics*, 131, 1057–1111.

KLEVEN, H., C. LANDAIS, AND J. E. SØGAARD (2018): "Children and Gender Inequality: Evidence from Denmark," *NBER Working Paper 24219*.

LEE, D. S. AND T. LEMIEUX (2010): "Regression Discontinuity Designs in Economics," *Journal of Economic Literature*, 2, 281–355.

LOCHNER, L. (2011): "Nonproduction Benefits of Education: Crime, Health, and Good Citizenship," *Handbook of the Economics of Education*, 4, 183–282.

LUNDBERG, S. (2018): "Report: Committee on the Status of Women in the Economics Profession (CSWEP)," *AEA Papers and Proceedings*, 108, 704–721.

MCCRARY, J. (2008): "Manipulation of the running variable in the regression discontinuity design: A density test," *Journal of econometrics*, 142, 698–714.

MULLIGAN, C. B. AND Y. RUBINSTEIN (2008): "Selection, Investment, and Women's Relative Wages over Time," *Quarterly Journal of Economics*, 123, 1061–1110.

NATIONAL SCIENCE FOUNDATION (2017): "Women, Minorities, and Persons with Disabilities in Science and Engineering," *National Science Foundation*.

PARKER, K. (2018): "Women in Majority-male Workplaces Report Higher Rates of Gender Discrimination," *Pew Research Center*, can be accessed at: http://www.pewresearch.org/fact-tank/2018/03/07/women-in-majority-male-workplaces-report-higher-rates-of-gender-discrimination/.

PSACHAROPOULOS, G. AND H. A. PATRINOS (2018): "Returns to Investment in Education: A Decennial Review of the Global Literature," *Education Economics*, 26, 445–458.

REUBEN, E. AND M. WISWALL (2016): "Preferences and Biases in Educational Choices and Labour Market Expectations: Shrinking the Black Box of Gender," *Economic Journal*, 127, 2153–2186.

SARSONS, H. (2017): "Gender Differences in Recognition for Group Work," *Working Papers*.

SVENSSON, L.-G. (1999): "Strategy-proof Allocation of Indivisible Goods," *Social Choice and Welfare*, 16, 557–567.

WISWALL, M. AND B. ZAFAR (2015): "Determinants of College Major Choice: Identification Using an Information Experiment," *Review of Economic Studies*, 82, 791–824.

WOLFERS, J. (2015): "Even Famous Female Ecnomists Get No Respect," *New York Times*, posted on November 11 and can be accessed at: https://www.nytimes.com/2018/02/02/business/why-womens-voices-are-scarce-in-economics.html.

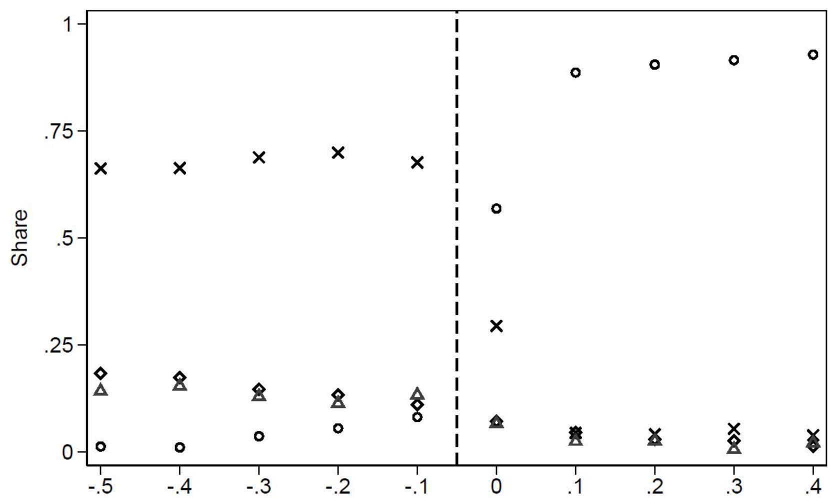
——— (2018): "Why Women's Voices Are Scarce in Economics," *New York Times*, can be accessed at: https://www.nytimes.com/2018/02/02/business/why-womens-voices-are-scarce-in-economics.html.

WU, A. H. (2018): "Gendered Language on the Economics Job Market Rumors Forum," *AEA Papers and Proceedings*, 108, 175–179.

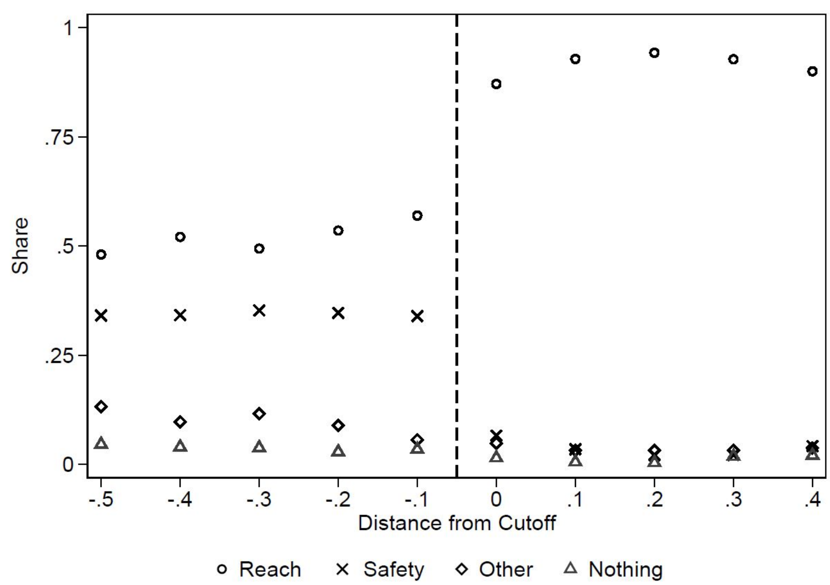
ÑOPO, H., N. DAZA, AND J. RAMOS (2011): "Gender Earnings Gaps in the World," *World Development Report 2012: Gender Equality and Development, Background Paper*.

7 Figures

Figure 1: Outcome of Applications



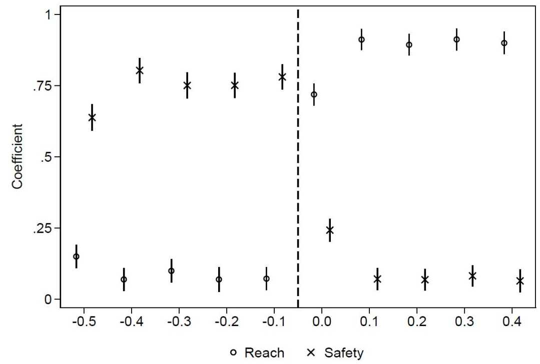
(a) Quota 1 Applications



(b) Quota 2 Applications

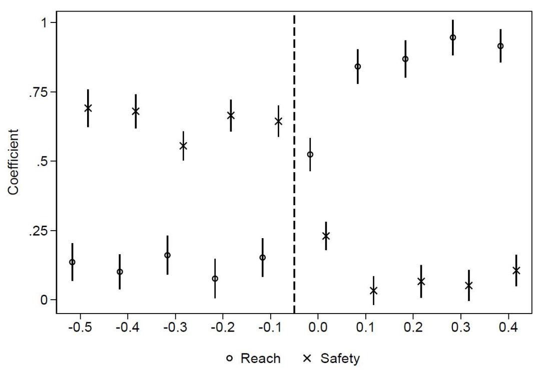
Note: This figure shows the different possible outcomes of an application to a program of study. In each panel we show the share of the sample of students admitted to their reach, safety, other, or no program by application score bins. We use the sample of 10,944 students who have a reach–safety programs pair in application years 1994 to 2006. We pool all programs and years and present the distance between the application score and the admission cutoff of the reach program on the x-axis. Panel (a) shows the subsample of applications made through the Quota 1 system of admission (6,658 observations) and Panel (b) shows Quota 2 applications (4,286 observations).

Figure 2: First-stage Results

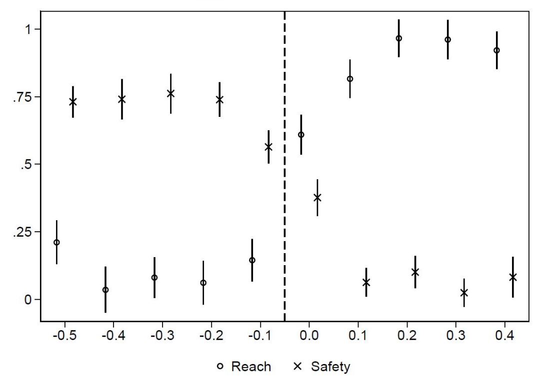


(a) Earnings: All

Gender Earnings Gap



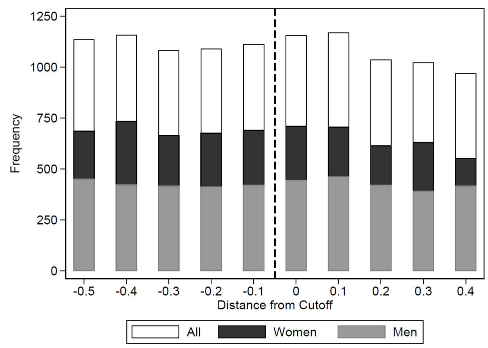
(b) Women



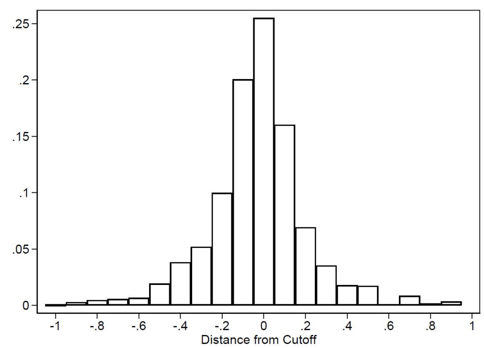
(c) Men

Note: This figure presents the graphical illustration of the first-stage earnings and earnings gap results. We use the sample of 6,658 students who applied through the Quota 1 system of admission and have a reach–safety programs pair in application years 1994 to 2006. We pool all programs and years and present the distance between the application score and the admission cutoff of the reach program on the x-axis. Results present the coefficients estimated separately for reach and safety programs using 10 dummy bins centered around the admission cutoff. Each student has exactly one dummy bin variable equal to 1, with the rest equal to zero. In Panel (a) the dependent variable is earnings of enrollees in the program of admission, and we plot the coefficients of earnings in the reach and safety programs interacted with distance from cutoff dummies. In Panels (b) and (c), the dependent variable is the gender earnings gap of enrollees in the program of admission and we plot the coefficients of gender earnings gaps in the reach and safety programs interacted with distance from cutoff dummies. Earnings are averaged 10 to 12 years after application. Point estimates are given along with the 95% confidence interval.

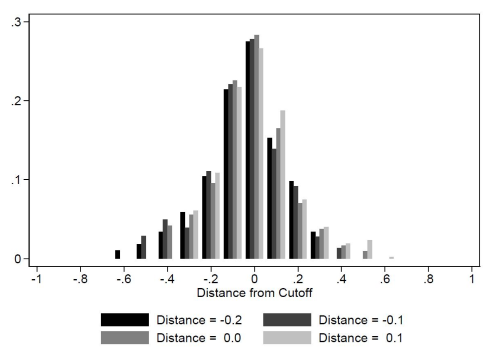
Figure 3: Bunching and Annual Changes in Cutoffs



(a) Total Counts



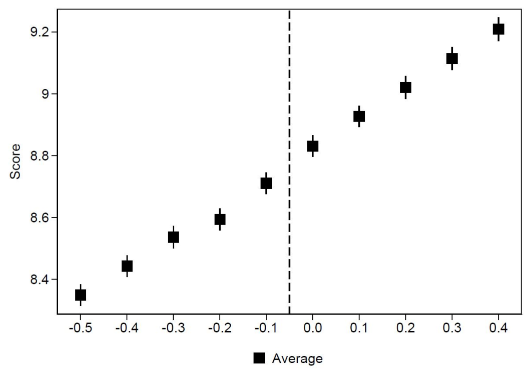
(b) Annual Changes Cutoff



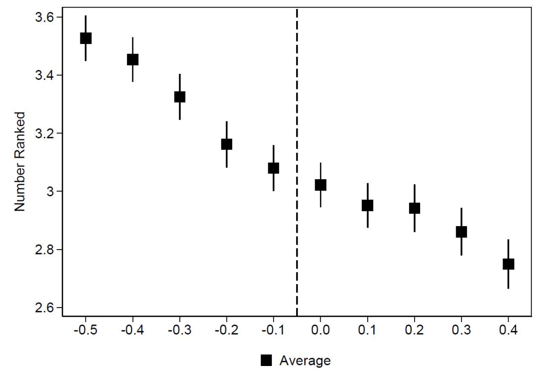
(c) Change in Cutoff Given Distance from Reach

Note: This figure presents the number of applications around admission cutoffs, and year-to-year changes in admission cutoffs. We use the sample of 10,944 students who have a reach–safety programs pair in application years 1994 to 2006. Panel (a) presents the number of applications around admission cutoffs for the entire sample, and for women and men separately. We pool all programs and years and present the distance between the application score and the admission cutoff of the reach program on the x-axis. Panel (b) presents the difference between the admission cutoff of the student's reach program in the year of application and the cutoff for the same program in the previous year. We restrict the histogram to changes in program admission cutoffs that are between –1 and 1 because they account for most of the data; 1,262 observations are for programs that have an increase in admission cutoff of greater than 1, and 29 observations are for programs that have a decrease in admission cutoff of less than –1. Panel (c) shows the year-to-year changes in the admission cutoff of a student's reach program conditional on the student's distance to the admission cutoff of the reach program in the year prior to his or her application.

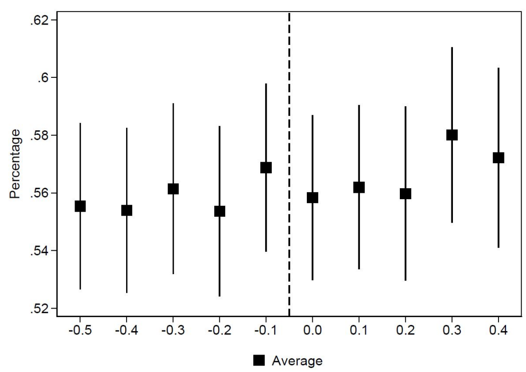
Figure 4: Covariate Balance



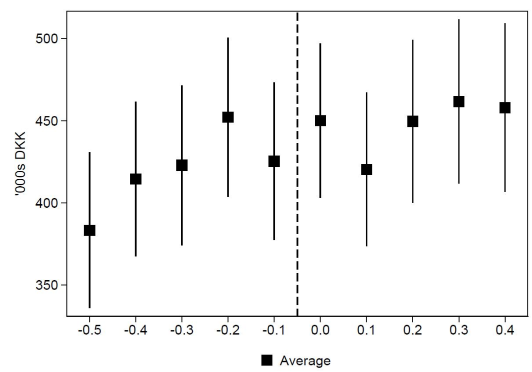
(a) High School Score



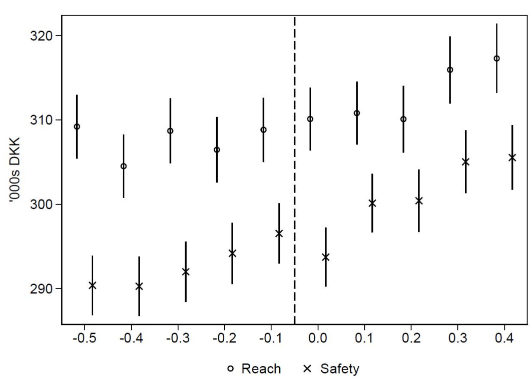
(b) Number of Programs Ranked



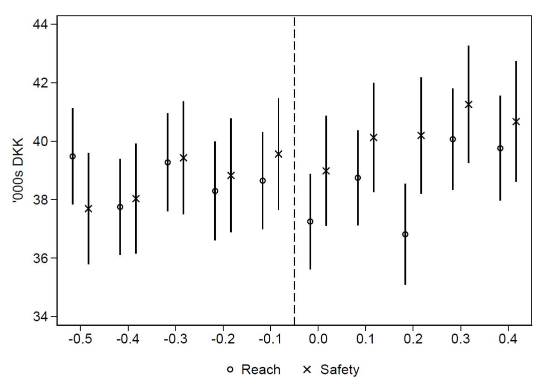
(c) Same Field for Reach and Safety



(d) Parental Wealth



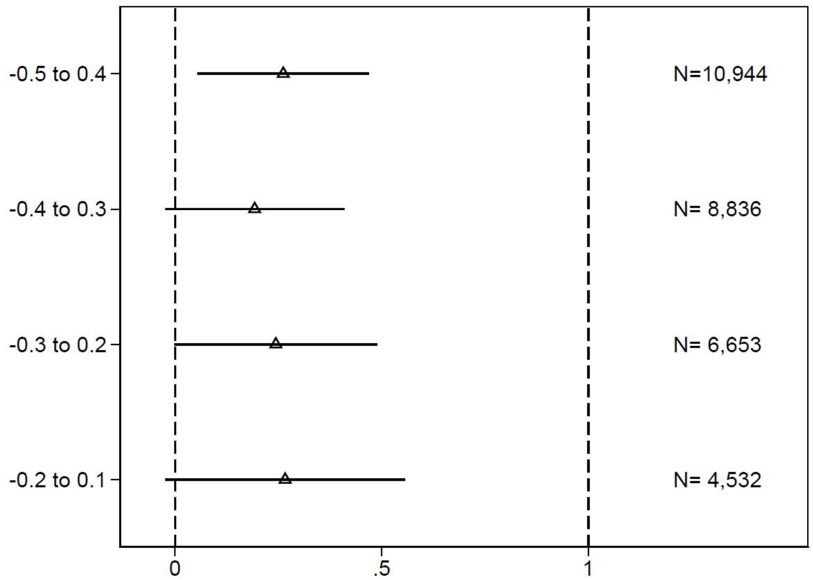
(e) Earnings



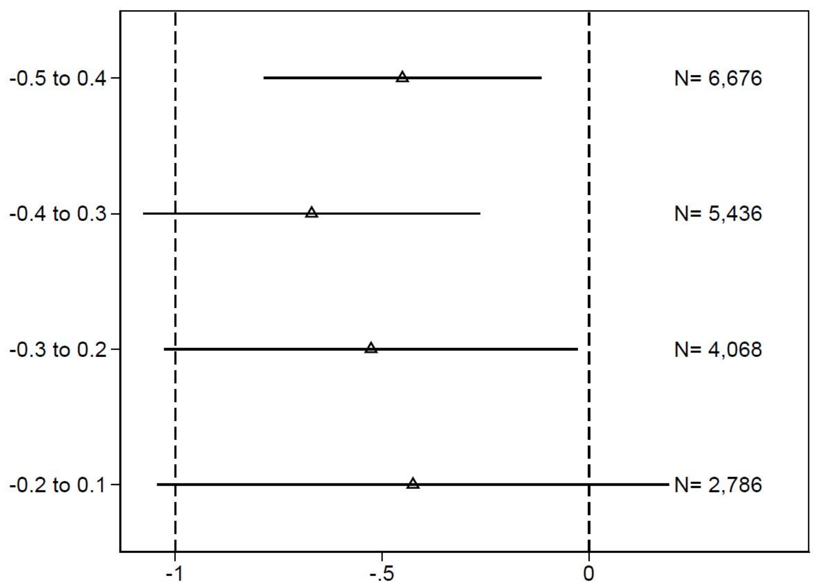
(f) Gender Earnings Gap

Note: This figure presents averages of the covariates used in the analysis by application score bins. We use the sample of 10,944 students who have a reach–safety programs pair in application years 1994 to 2006. We pool all programs and years and present the distance between the application score and the admission cut-off of the reach program on the x-axis. Point estimates are given along with the 95% confidence interval.

Figure 5: Alternative Bandwidths



(a) Earnings



(b) Gender Earnings Gap

Note: This figure presents the robustness of our main results to the selection of different bandwidths. Panel (a) uses the sample of 10,944 students who have a reach–safety programs pair in application years 1994 to 2006 and shows the estimated coefficient of δ from equation (2) for earnings, setting $\overline{g_{p_i}}$ to 0. Panel (b) uses the subsample of 6,676 women who have a reach–safety programs pair in application years 1994 to 2006 and shows the estimated coefficient of λ on the gender earnings gap from equation (2). A coefficient of 0 represents a model in which individual outcomes reflect only selection, and a coefficient of 1 or –1 represents a model in which individual outcomes reflect only causal effects. Point estimates and 95% confidence intervals are shown with the number of observations included in the different bandwidth ranges.

8 Tables

Table 1: Descriptive Statistics

	All applications (320,308 obs.)		Used sample (10,944 obs.)	
	Mean	Std. dev	Mean	Std. dev
A. Demographics				
Age	21.4	3.7	21.1	2.9
Male percent	37.3	–	39.0	–
Parental income (000's DKK)	298.1	153.2	328.6	168.2
Parental wealth (000's DKK)	348.7	714.7	432.9	815.4
Parental education (Years)	12.7	3.2	13.5	3.2
B. Application				
Grade	8.3	1.0	8.8	0.7
Nbr programs ranked	2.2	1.4	3.1	1.4
Accepted into reach	–	–	0.6	0.5
Rank of reach	–	–	1.2	0.6
Rank of safety	–	–	2.3	0.7
Rank of enrollment prog.	–	–	1.6	0.9
C. Reach-Safety pairs				
Grade distance from the cutoff	–	–	−0.1	0.3
Same field			0.6	–
Same institution	–	–	0.3	–
D. Earnings				
All	300.4	154.5	306.9	165.0
Women	273.9	134.5	286.6	149.2
Men	345.1	174.5	338.6	182.7
Enrollment cohort average	294.1	58.3	303.6	63.7
Enrollment cohort gender gap	46.5	34.1	40.2	29.0
Reach-Safety gap	–	–	16.0	59.8
Reach-Safety gap in gender gap	–	–	−1.1	37.7

Note: This table presents descriptive statistics for the population of students applying to university for the first time in application years 1994 to 2006 (all applications; left), as well as the sample of 10,944 students who have a reach–safety programs pair in these application years (used sample; right). Panel A presents demographics. Parental income, wealth and education represent the average values of both parents (or the values for a single parent). Panel B presents information typically found in a student's application. The student's score summarizes their high school grades. The number of priorities ranked refers to the number of programs to which the student applied. For the sample of students with a reach–safety pair, we present the rankings of the reach, safety and enrollment programs. Panel C describes the reach–safety pairs. The distance from the cutoff represents the difference between the student's grade and the cutoff of the reach program. Panel D presents information on earnings. Earnings are averaged 10 to 12 years after the application to the university program. We show the average earnings for all students and for the subsamples of women and men. The student's enrollment cohort earnings average is the average earnings of the cohort of students admitted to the same program–university major. The enrollment cohort earnings gender gap is the difference between the earnings of men and women for the cohort of students admitted to the same program–university major. The reach–safety earnings gap is the difference between the earnings of students admitted to their reach program and the earnings of students admitted to their safety program. The reach–safety gender gap is the difference between the earnings gender gap for students admitted to their reach program and the earnings gender gap for students admitted to their safety program.

Table 2: Fields of Study

	All applications (320,308 obs.)				Used sample (10,944 obs.)			
	Enrollment Cohort Averages (000s DKK)				Women (11,725 obs.)		Men (7,473 obs.)	
	Average Earnings (s.d.)	Gender Gap (s.d.)	Gender Gap / Average Earnings	Share of Women	Reach	Safety	Reach	Safety
Business	361	75	21 %	45 %	6.9 %	8.2 %	14.5 %	15.2 %
	(49)	(32)						
Education	289	33	12 %	69 %	10.3 %	10.8 %	5.6 %	7.3 %
	(29)	(14)						
Engineering	358	61	17 %	32 %	0.3 %	1.2 %	1.7 %	2.8 %
	(72)	(61)						
Health	276	41	15 %	85 %	23.2 %	22.5 %	13.0 %	7.5 %
	(53)	(29)						
Humanities	241	29	12 %	65 %	21.0 %	26.8 %	19.5 %	23.1 %
	(33)	(36)						
Law	379	51	13 %	58 %	5.1 %	2.2 %	6.5 %	2.8 %
	(11)	(7)						
Basic Programs	288	67	23 %	63 %	0.0 %	1.7 %	0.0 %	1.8 %
	(18)	(21)						
Science	309	32	10 %	47 %	6.4 %	9.2 %	6.2 %	11.7 %
	(43)	(26)						
Social Science	319	34	11 %	53 %	21.1 %	13.0 %	25.7 %	19.7 %
	(47)	(24)						
Technology	326	63	19 %	30 %	5.7 %	4.5 %	7.4 %	8.2 %
	(51)	(50)						
Total					100 %	100 %	100 %	100 %

Note: This table presents descriptive statistics by fields of study for the population of students applying to university for the first time in application years 1994 to 2006 (all applications; left), as well as the sample of 10,944 students who have a reach–safety programs pair in these application years (used sample; right). The left panel shows the average earnings and gender earnings gap in each field, as well as the gender earnings gap as a percentage of the average earnings and the share of women enrolled in each fields. Earnings are averaged 10 to 12 years after the application to the university program. The standard errors show the variation over programs within a broad field. The right panel shows the distribution of fields as reaches and safeties for men and women separately.

Table 3: Density Manipulation Tests

Bandwitdh:	-0.5 to 0.4		-0.4 to 0.3		-0.3 to 0.2	
Model:	Restricted	Unrestricted	Restricted	Unrestricted	Restricted	Unrestricted
A. Full Sample						
Conventional S.E.	0.344	0.746	0.959	0.499	0.583	0.223
Jackknife S.E.	0.317	0.753	0.956	0.519	0.540	0.266
B. Only Women						
Conventional S.E.	0.094	0.222	0.416	0.800	0.843	0.786
Jackknife S.E.	0.077	0.235	0.380	0.809	0.825	0.804
C. Only Men						
Conventional S.E.	0.553	0.307	0.347	0.160	0.256	0.105
Jackknife S.E.	0.530	0.319	0.312	0.180	0.205	0.137

Note: This table presents the p-values for the hypothesis test in equation (8) under different assumptions about the data generating process. Panel A. uses the entire sample, while Panels B and C restrict the sample to women and men, respectively. We present the results of the test using local-linear polynomials for bandwidths varying from +/- 0.5 to +/- 0.3 around the admission cutoff. The unrestricted version of the test allows both estimators $\hat{f}_{+,p}(h)$ and $\hat{f}_{-,p}(h)$ to be unrelated. The restricted version of the test is more powerful but assumes that the c.d.f. and higher-order derivatives of the running variable are equal for treated and untreated groups at the cutoff even when $f_{-} \neq f_{+}$. We show the results using both conventional and jackknife standard errors.

Table 4: Gender Earnings Gap

		Gender earnings gap		Gender-specific earnings	
	All	Women	Men	Women	Men
A. Instrumental variable estimation					
Program earnings (All)	0.26	0.48	0.06		
	(0.11)	(0.13)	(0.19)		
Within-program gender earnings gap		-0.45	-0.09		
		(0.18)	(0.25)		
Program earnings (Women)				0.76	0.25
				(0.26)	(0.36)
Program earnings (Men)				-0.31	-0.01
				(0.23)	(0.29)
B. OLS benchmark					
Program earnings (All)	0.80	0.79	0.75		
	(0.02)	(0.03)	(0.05)		
Within-program gender earnings gap		-0.14	0.42		
		(0.07)	(0.10)		
Program earnings (Women)				0.69	-0.02
				(0.08)	(0.11)
Program earnings (Men)				0.14	0.85
				(0.06)	(0.09)
N	10,944	6,676	4,268	6,676	4,268

Note: This table presents the results of estimating equation (2) by IVs (Panel A) and by OLS (Panel B) for the gender earnings gap and for the subsamples of women and men. Each column presents the second stage of an IV regression. We use the sample of 10,944 students who have a reach–safety programs pair in application years 1994 to 2006 and present the results for a bandwidth of +/− 0.5 around the admission cutoff of the reach program. The outcome variable for the regressions is the student's subsequent earnings averaged 10 to 12 years after admission. The first column sets $\overline{g_{p_i}}$ to 0. The second and third columns separate the sample based on gender and use the average program earnings of enrollees and the earnings gap between men and women for the program of enrollment. The fourth and fifth columns separate the sample based on gender and use the average program earnings of men and women separately. The IV estimation controls for distance from cutoff fixed effects (bin FE), average earnings of the reach program, average earnings of the safety program, calendar year FE, the probability of being accepted given last year's admission cutoff interacted with the average earnings of the reach and safety programs. The OLS results control for parents' wealth, parents' income and parents' education, student grade, age and gender, and calendar year FE.

Table 5: Gender Earnings Gap Channels

	Specifications				
	(1)	(2)	(3)	(4)	(5)
Program earnings (All)	0.48	0.48	0.49	0.46	0.48
	(0.13)	(0.13)	(0.13)	(0.13)	(0.11)
Within-program gender earnings gap	-0.45	-0.45	-0.42	-0.45	-0.35
	(0.18)	(0.18)	(0.18)	(0.18)	(0.16)
Married		-1.98	12.30	11.31	7.04
		(3.44)	(4.07)	(4.04)	(3.62)
Nb. of kids			-17.55	-18.87	-25.39
			(4.47)	(4.43)	(3.97)
Graduated				38.94	31.08
				(3.46)	(3.10)
Any labor income					282.82
					(7.02)

Note: This table presents the results of estimating equation (2) by IVs for the subsample of women. Each column presents the second stage of an IV regression for which we vary the control variables, all measured 10 to 12 years after application and all scaled in thousands. The first column repeats the baseline result. The second column controls for a dummy variable indicating marital situation, the third column adds number of children fixed effects, the fourth column indicates whether the student has graduated from the reach program, and the fifth column adds a dummy variable for no labor income earnings. We use the sample of 6,676 women who have a reach–safety programs pair in application years 1994 to 2006 and present the results for a bandwidth of +/− 0.5 around the admission cutoff of the reach program. The outcome variable for the regressions is the student's subsequent earnings averaged 10 to 12 years after admission.

Table 6: Gender Earnings Gap Dynamics

	Years after enrollment		
	7 to 9	10 to 12	13 to 15
A. Instrumental variable estimation			
Program earnings (All)	0.44	0.48	0.54
	(0.13)	(0.13)	(0.17)
Within-program	-0.28	-0.45	-0.80
gender earnings gap	(0.16)	(0.18)	(0.23)
B. OLS Results			
Program earnings (All)	0.86	0.79	0.72
	(0.03)	(0.03)	(0.04)
Within-program	-0.15	-0.14	-0.22
gender earnings gap	(0.07)	(0.07)	(0.08)

Note: This table presents the results of estimating equation (2) by IVs (Panel A) and by OLS (Panel B). In Panel A, each column presents the second stage of an IV regression. We use the sample of 6,676 women who have a reach–safety programs pair in application years 1994 to 2006 and present the results for a bandwidth of +/− 0.5 around the admission cutoff of the reach program. The outcome variable for the regressions is the student's subsequent earnings averaged seven to nine years, 10 to 12 years and 13 to 15 years after admission in columns 1, 2 and 3, respectively. The second column repeats the baseline result. The OLS results control for parents' wealth, parents' income and parents' education, student grade, age and gender, and calendar year FE.

Table 7: University and Program Decomposition

	Women	Men	All	Same university reach and safety
Program earnings (All)	0.50	0.01		0.13
	(0.13)	(0.20)		(0.16)
University gender earnings gap	-0.76	0.22		
	(0.30)	(0.49)		
Program gender earnings gap	-0.42	-0.12		
minus university gender earnings gap	(0.18)	(0.25)		
University earnings			0.57	
			(0.16)	
Program earnings			0.17	
minus university earnings			(0.11)	
N	6,676	4,268	10,944	3,696

Note: This table presents the results of estimating equation (2) by IVs. Each column presents the second stage of an IV regression. The first column repeats the baseline results. The second column restricts the sample used to students choosing the same university for their reach and safety programs and sets $\overline{g_{p_i}}$ to 0. The third column uses the full reach–safety programs pair sample, but decomposes the average program earnings into average university earnings and the difference between average program earnings and average university earnings. The fourth and fifth columns restrict the sample used to women and men, respectively, and decomposes the gender earnings gap into the average gender earnings gap at the university level and the difference between the average university gender earnings gap and the average program gender earnings gap. We use the sample of 10,944 students who have a reach–safety programs pair in application years 1994 to 2006 and present the results for a bandwidth of +/− 0.5 around the admission cutoff of the reach program. The outcome variable for the regressions is the student's subsequent earnings averaged 10 to 12 years after admission.

Table 8: Field of Study and Program Decomposition

	Women	Men	All	Same field reach and safety
Program earnings (All)	0.48	0.10		0.38
	(0.13)	(0.20)		(0.18)
Field gender earnings gap	-0.34	-0.23		
	(0.22)	(0.37)		
Program gender earnings gap minus field gender earnings gap	-0.55	0.03		
	(0.24)	(0.29)		
Field earnings			0.25	
			(0.11)	
Program earnings minus field earnings			0.39	
			(0.19)	
N	6,676	4,268	10,944	6,153

Note: This table presents the results of estimating equation (2) by IVs for earnings and earnings gender gaps. Each column presents the second stage of an IV regression. The first column repeats the baseline results. The second column restricts the sample used to students choosing the same field for their reach and safety programs and sets $\overline{g_{p_i}}$ to 0. The third column uses the full reach–safety programs pair sample, but decomposes the average program earnings into average field earnings and the difference between average program earnings and average field earnings. The fourth and fifth columns restrict the sample used to women and men, respectively, and decomposes the gender earnings gap into the average gender earnings gap at the field-of-study level and the difference between the average field gender earnings gap and the average program gender earnings gap. We use the sample of 10,944 students who have a reach–safety programs pair in application years 1994 to 2006 and present the results for a bandwidth of +/− 0.5 around the admission cutoff of the reach program. The outcome variable for the regressions is the student's subsequent earnings averaged 10 to 12 years after admission.

Table 9: Share of Men in the Program

	Women (Baseline)	Women	Men
Program earnings (All)	0.48	0.51	0.22
	(0.09)	(0.13)	(0.19)
Within-program	-0.27	-0.45	-0.03
gender earnings gap	(0.13)	(0.18)	(0.26)
Share of men in program		-21.87	-118.32
		(32.46)	(44.56)
N	10,944	6,676	4,268

Note: This table presents the results of estimating equation (2) by IVs for earnings and earnings gender gaps. Each column presents the second stage of an IV regression. The first column repeats the baseline results. The second and third columns restrict the sample used to women and men, respectively, and control for the share of males in the program of admission. We use the sample of 10,944 students who have a reach–safety programs pair in application years 1994 to 2006, and present the results for a bandwidth of +/− 0.5 around the admission cutoff of the reach program. The outcome variable for the regressions is the student's subsequent earnings averaged 10 to 12 years after admission.

Table 10: Robustness to Treatment Results

	All		Women		Men	
	IV	OLS	IV	OLS	IV	OLS
A. Initial program; baseline						
Program earnings (All)	0.26	0.8	0.48	0.79	0.06	0.75
	(0.11)	(0.02)	(0.13)	(0.03)	(0.19)	(0.05)
Within-program			-0.45	-0.14	-0.086	0.42
gender earnings gap			(0.18)	(0.06)	(0.25)	(0.10)
B. Initial to 4th year programs						
Program earnings (All)	0.46	0.94	0.7	0.97	0.19	1.03
	(0.16)	(0.02)	(0.21)	(0.03)	(0.28)	(0.05)
Within-program			-0.57	-0.16	-0.056	0.15
gender earnings gap			(0.26)	(0.06)	(0.38)	(0.09)
C. 4th year program						
Program earnings (All)	0.6	0.94	0.8	0.91	0.34	0.99
	(0.19)	(0.02)	(0.24)	(0.03)	(0.33)	(0.05)
Within-program			-0.61	-0.16	0.051	0.08
gender earnings gap			(0.29)	(0.06)	(0.42)	(0.09)

Note: This table presents the results of estimating equation (2) by IVs and by OLS for the full sample and the subsamples of women and men. We use the sample of 10,944 students who have a reach–safety programs pair in application years 1994 to 2006 and present the results for a bandwidth of +/− 0.5 around the admission cutoff of the reach program. The earnings are averaged 10 to 12 years after the application to university programs. The outcome variable for the first-stage regression is a measure of the average earnings of students admitted to the same program, and the outcome variable for the other regressions is the student's subsequent earnings. Panel A defines program earnings as the average subsequent earnings of students who initially enroll in that program; Panel B defines program earnings as the average subsequent earnings across all programs studied by that student over their first four years at university; and Panel C defines program earnings as the average subsequent earnings of the program the student is enrolled in four years after initial admission. The IV, first-stage and reduced-form estimates control for distance from cutoff fixed effects (bin FE), average earnings of the reach program, average earnings of the safety program, calendar year FE, the probability of being accepted given last year's admission cutoff interacted with the average earnings of the reach and safety programs. The OLS equation controls for parents' wealth, parents' income and parents' education, student grade, age and gender, and calendar year FE.

Table 11: Robustness to Specifications

	All		Women		Men	
	IV	OLS	IV	OLS	IV	OLS
A. Baseline						
Program earnings (All)	0.26	0.8	0.48	0.79	0.06	0.75
	(0.11)	(0.02)	(0.13)	(0.03)	(0.19)	(0.05)
Within-program			-0.45	-0.14	-0.086	0.42
gender earnings gap			(0.18)	(0.06)	(0.25)	(0.10)
B. Baseline with reach and safety FE						
Program earnings (All)	0.31	0.8	0.64	0.79	0.014	0.79
	(0.11)	(0.02)	(0.14)	(0.03)	(0.19)	(0.05)
Within-program			-0.36	-0.14	0.085	-0.14
gender earnings gap			(0.20)	(0.06)	(0.28)	(0.10)
C. Baseline without lagged instruments						
Program earnings (All)	0.25	0.8	0.43	0.79	0.049	0.75
	(0.08)	(0.02)	(0.10)	(0.03)	(0.15)	(0.05)
Within-program			-0.46	-0.14	-0.081	0.42
gender earnings gap			(0.18)	(0.06)	(0.25)	(0.09)
D. Baseline with cutoffs and score						
Program earnings (All)	0.27	0.8	0.5	0.79	0.078	0.75
	(0.11)	(0.02)	(0.13)	(0.03)	(0.19)	(0.05)
Within-program			-0.46	-0.14	-0.11	0.42
gender earnings gap			(0.18)	(0.06)	(0.25)	(0.09)
E. Baseline with cutoffs, score and demographics						
Program earnings (All)	0.28	0.8	0.5	0.79	0.097	0.75
	(0.10)	(0.02)	(0.13)	(0.03)	(0.19)	(0.05)
Within-program			-0.47	-0.14	-0.13	0.42
gender earnings gap			(0.18)	(0.06)	(0.25)	(0.09)

Note: This table presents the results of estimating equation (2) by IVs and by OLS for the full sample and the subsamples of women and men. We use the sample of 10,944 students who have a reach–safety programs pair in application years 1994 to 2006 and present the results for a bandwidth of +/− 0.5 around the admission cutoff of the reach program. The earnings are averaged 10 to 12 years after the application to university programs. The outcome variable for the first-stage regression is the average subsequent earnings of students who initially enroll in that program, and the outcome variable for the other regressions is the student's subsequent earnings. We vary the controls used in the estimations in Panels A, B, C, D and E. Panel A presents our baseline specification and controls for distance from cutoff fixed effects (bin FE), average earnings of the reach program, average earnings of the safety program, calendar year FE and the probability of being accepted given last year's admission cutoff interacted with the average earnings of the reach and safety programs. Panel B adds reach and safety programs fixed effects. Panel C omits the probability of being accepted given last year's admission cutoff interacted with the average earnings of the reach and safety programs. Panel D adds the reach and the safety program cutoffs and the student's score. Panel E adds demographics (parents' wealth, parents' income and parents' education, and student's age and gender).

Table 12: Robustness to Income-program Measures

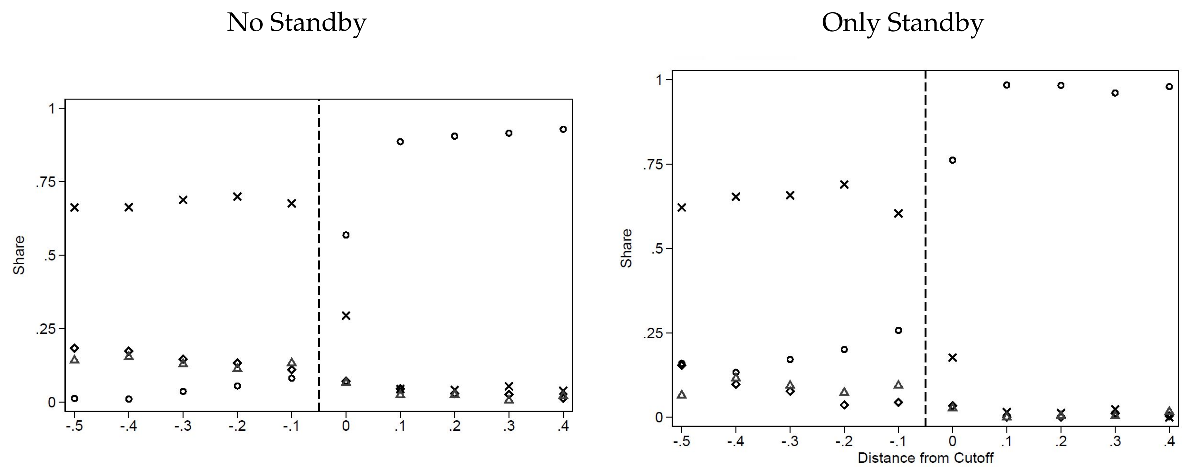
	Excluding own cohort		
	All	Women	Men
A. Instrumental variable estimation			
Program earnings (All)	0.21	0.40	0.05
	(0.09)	(0.13)	(0.19)
Within-program		-0.25	-0.18
gender earnings gap		(0.16)	(0.20)
B. OLS benchmark			
Program earnings (All)	0.78	0.74	0.74
	(0.03)	(0.04)	(0.05)
Within-program		-0.02	0.31
gender earnings gap		(0.09)	(0.09)

Note: This table presents the results of estimating equation (2) by IVs (Panel A) and by OLS (Panel B) for the gender earnings gap and for the subsamples of women and men. Each column presents the second stage of an IV regression. We use the sample of 10,944 students who have a reach–safety programs pair in application years 1994 to 2006 and present the results for a bandwidth of +/− 0.5 around the admission cutoff of the reach program. The outcome variable for the regressions is the student's subsequent earnings averaged 10 to 12 years after admission. We present an alternative way of calculating the income averages for the reach, safety and enrollment programs. We calculate these averages omitting the cohort-year peers of the enrollee, and therefore omits cohort-year peers of the enrollee and future cohorts from the average calculation. The first column repeats the baseline results with these alternative definitions of program earnings. The second and third columns separate the sample based on gender and use these alternative definitions of average program earnings and the earnings gap between males and females. The OLS results control for parents' wealth, parents' income and parents' education, student grade, age and gender, and calendar year FE.

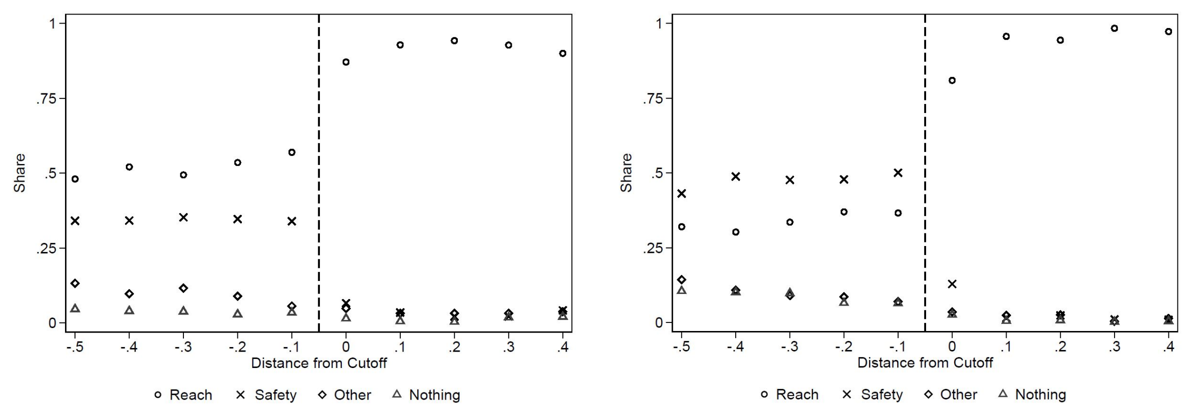
A Appendix

A.1 Additional Figures

Figure A1: Outcome of Applications



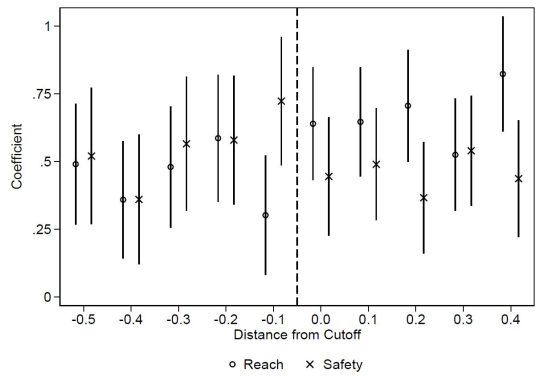
(a) Quota 1



(b) Quota 2

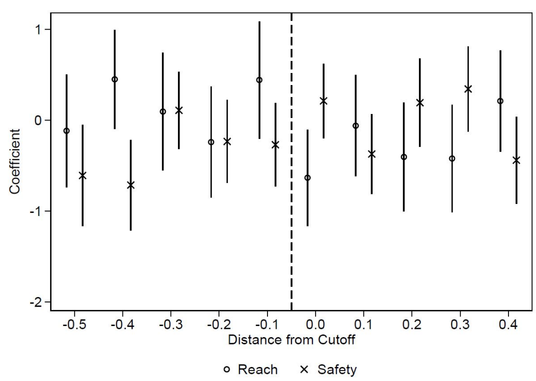
Note: This figure shows the different possible outcomes of an application to a program of study. In each panel we show the share of the sample of students admitted to their reach, safety, other, or no program by application score bins. We use the sample of 10,944 students who have a reach–safety programs pair in application years 1994 to 2006. We pool all programs and years and present the distance between the application score and the admission cutoff of the reach program on the x-axis. Panel (a) shows the subsample of applications made through the Quota 1 system of admission (6,658 observations) and Panel (b) shows Quota 2 applications (4,286 observations).

Figure A2: Quota 1 Applications: Reduced-form Results

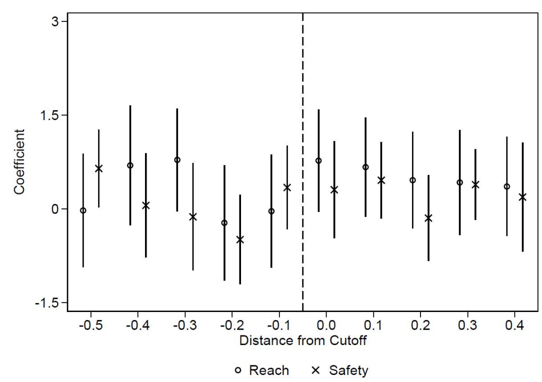


(a) Earnings: All

Gender Earnings Gap



(b) Women

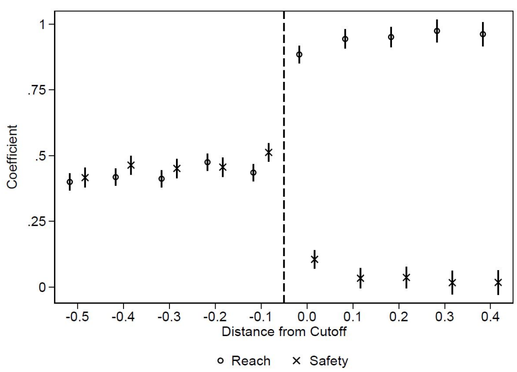


(c) Men

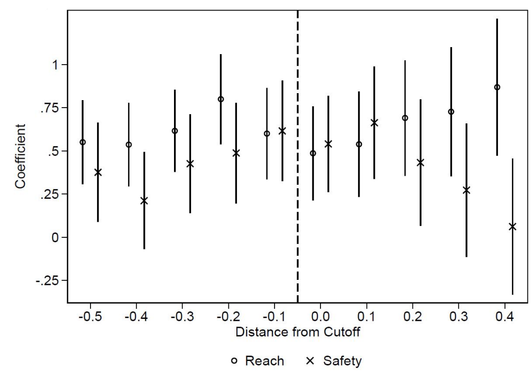
Note: This figure presents the graphical illustration of the reduced-form results for earnings and the gender earnings gap. We use the sample of 6,658 students who applied through the Quota 1 system of admission and who have a reach–safety programs pair in application years 1994 to 2006. We pool all programs and years together and present the distance between the application score and the admission cutoff of the reach program on the x-axis. Results present the coefficients estimated separately for reach and safety programs using 10 dummy bins centered around the admission cutoff. Each student has exactly one dummy bin variable equal to 1, with the rest equal to zero. Panel (a) shows the results for the overall sample, Panel (b) show the results for the subsample of women (3,977 observations) and Panel (b) shows the subsample of men (2,662 observations). In the first-stage results, the dependent variable is earnings of enrollees in the program of admission and we plot the coefficients of earnings in the reach and safety programs interacted with distance from cutoff dummies. Earnings are averaged 10 to 12 years after application. Point estimates are given along with the 95% confidence interval.

Figure A3: Quota 2 Applications: First-stage and Reduced-form Results

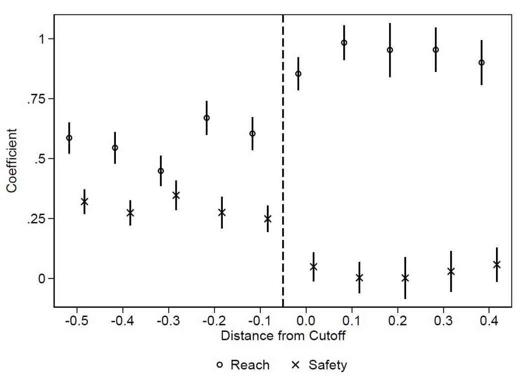
First-Stages | Reduced Forms



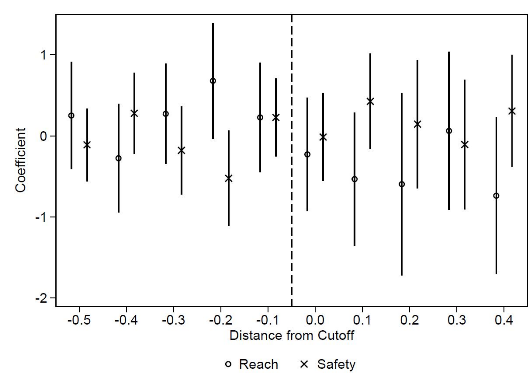
(a) Earnings: All



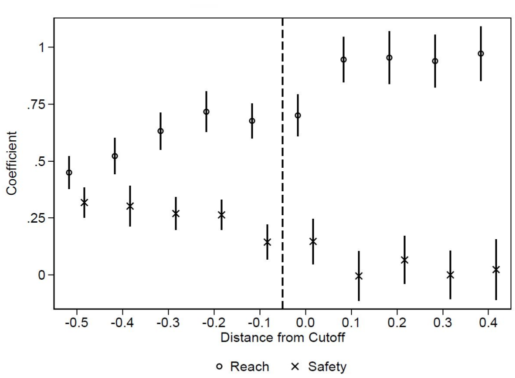
(b) Earnings: All



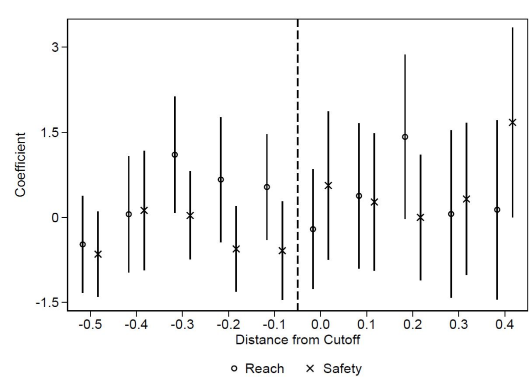
(c) Gender Earnings Gap: Women



(d) Gender Earnings Gap: Women



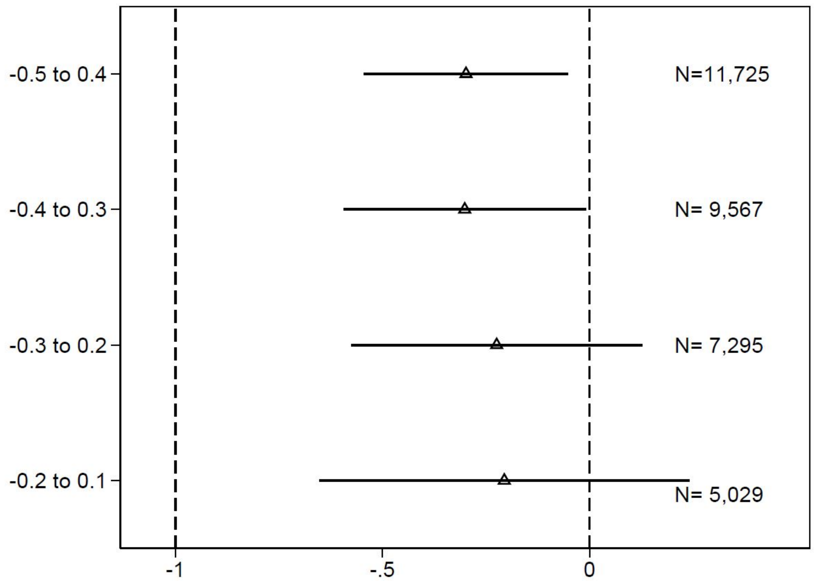
(e) Gender Earnings Gap: Men



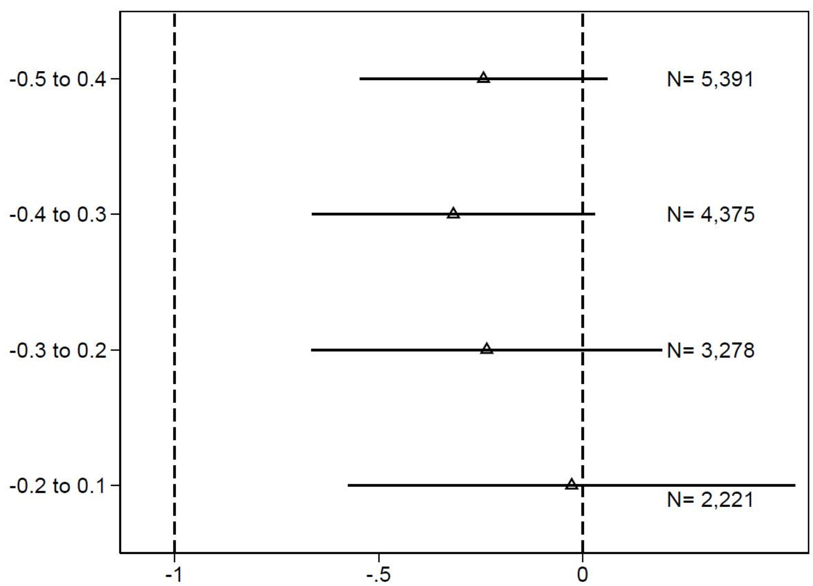
(f) Gender Earnings Gap: Men

Note: This figure presents a graphical illustration of the first-stage (left panel) and reduced-form (right panel) results for the gender earnings gap. We use the sample of 4,286 students who applied through the Quota 2 system and have a reach–safety programs pair in application years 1994 to 2006. We pool all programs and years and present the distance between the application score and the admission cutoff of the reach program on the x-axis. Results present the coefficients estimated separately for reach and safety programs using 10 dummy bins centered around the admission cutoff. Each student has exactly one dummy bin variable equal to 1, with the rest equal to zero. Panel (a) shows the results for earnings and Panels (b) and (c) show the results for the gender earnings gap for the subsamples of women (2,684 observations) and men (1,593 observations), respectively. In the first-stage results, the dependent variable is either earnings of enrollees in the program of admission or the gender earnings gap. We plot the coefficients of earnings and gender earnings gaps in the reach and safety programs interacted with distance from the cutoff dummies. In the reduced-form results, the dependent variable is the individual's earnings and we plot the coefficients of earnings and gender earnings gaps in the reach and safety programs interacted with distance from the cutoff dummies. Earnings are averaged 10 to 12 years after application. Point estimates are given along with the 95% confidence interval.

Figure A4: Alternative Bandwidths Robustness



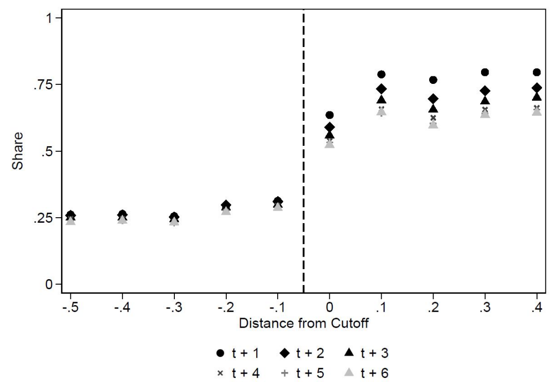
(a) All Admission Systems



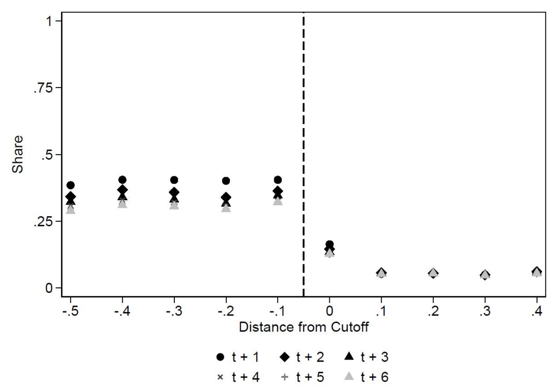
(b) Only Quota 1 (Including Standby)

Note: This figure presents the robustness of our main results to the selection of different bandwidths using two different sample of prospective students. Both panels use the subsample of women who have a reach–safety programs pair in application years 1994 to 2006 and shows the estimated coefficient of λ on the gender earnings gap from equation (2). Panel (a) uses the sample of 11,725 female students who applied on any of the admission system (either Quota 1 or Quota 2, and both with and without standby). Panel (b) uses the subsample of 5,391 female students who applied on Quota 1 both with and without standby. Point estimates and 95% confidence intervals are shown with the number of observations included in the different bandwidth ranges.

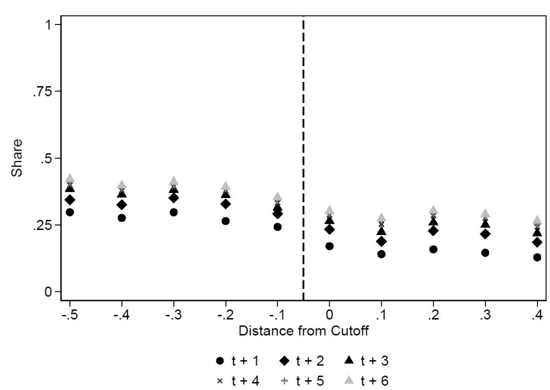
Figure A5: Proportion of Students in Reach, Safety, or Other Programs Over Time



(a) Reach



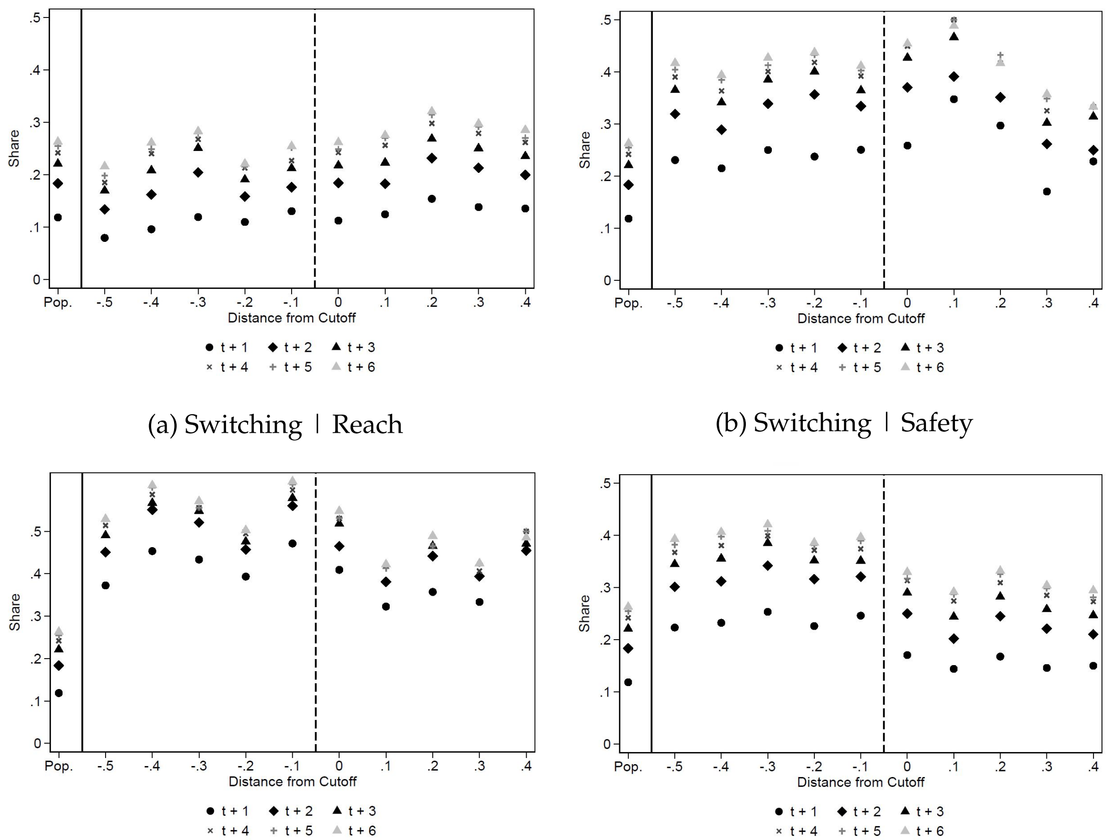
(b) Safety



(c) Other

Note: This figure presents the proportion of students whose latest admission data show that they are enrolled in their (a) reach, (b) safety, or (c) other program one to six years after their initial admission by application score bins. We pool all programs and present the distance between the application score and the admission cutoff of the reach program on the x-axis.

Figure A6: Switching Conditional on Reach, Safety or Other



(a) Switching | Reach

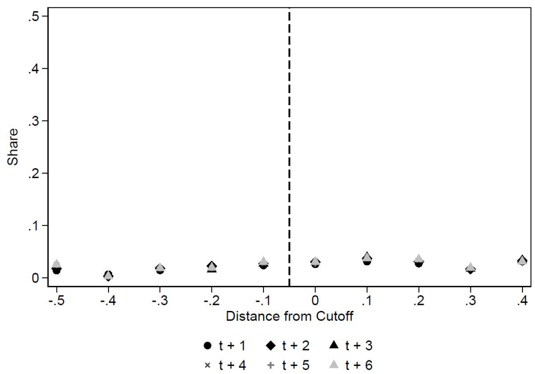
(b) Switching | Safety

(c) Switching | Other

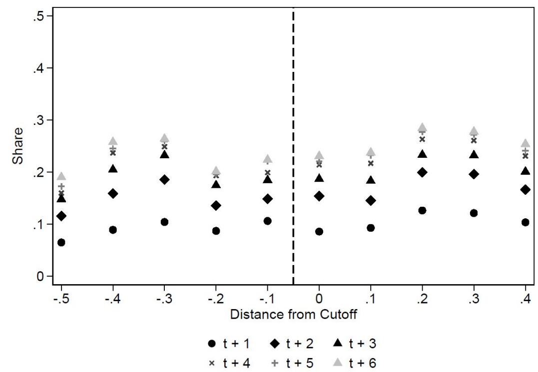
(d) Switching

Note: This figure presents the proportion of students whose latest admission data show that they have switched to another program given their initial admission into (a) their reach program, (b) their safety program, (c) another program, or (d) unconditionally one to six years after their initial admission by application score bins. We pool all programs and present the distance between the application score and the admission cutoff of the reach program on the x-axis.

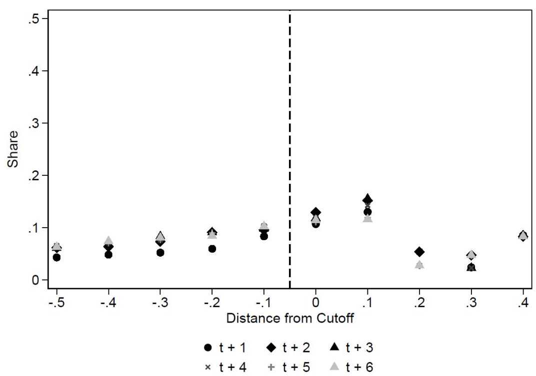
Figure A7: Transitions in and out of Reach, Safety or Other



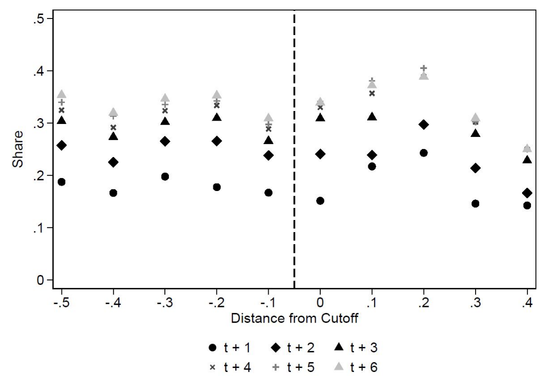
(a) Safety | Reach



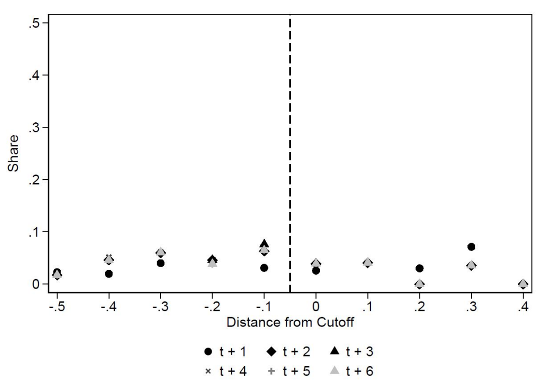
(b) Other | Reach



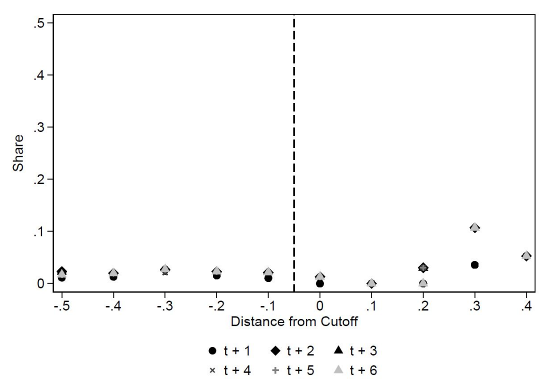
(c) Reach | Safety



(d) Other | Safety



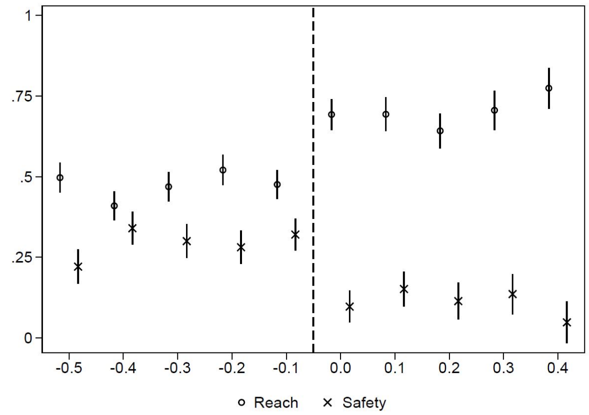
(e) Reach | Other



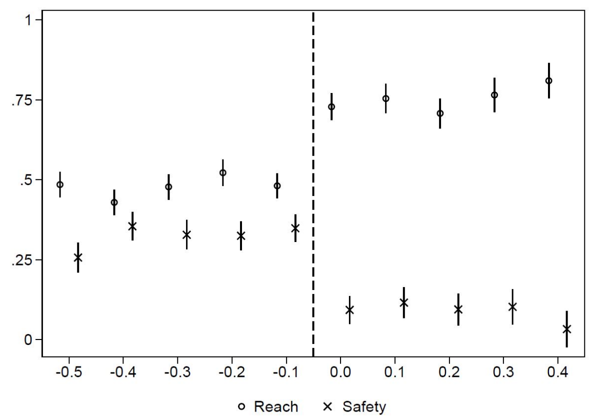
(f) Safety | Other

Note: This figure presents the proportion of students whose latest admission data show that they have transitioned (a) from their reach to their safety program, (b) from their reach program to another program, (c) from their safety to their reach program, (d) from their safety program to another program, (e) from another program to their reach program, or (f) from another program to their safety program one to six years after their initial admission by application score bins. We pool all programs and present the distance between the application score and the admission cutoff of the reach program on the x-axis.

Figure A8: Alternative First-stage Results



(a) Using study at year 4



(b) Using average of 4 first years of study

Note: This figure presents graphical representation of the first-stage results for alternative assignments of students to programs. The results show the impact on the average earnings of the program the student enrolls in as a function of the average earnings of their reach and safety programs, respectively, for each distance from the cutoff bin. We average earnings for 10 to 12 years after application for all students admitted to a given program. Panel (a) averages this measure over the first four years of the program to which a student has been admitted. For example, a student studying business for two years and health for two years would be attributed the average of future business and health earnings. Panel (b) attributes the earnings of the program in which the student was enrolled four years after application. Point estimates for the reach and safety programs are given, along with the 95% confidence interval.

A.2 Additional Tables

Table A1: Descriptive Statistics: Demographics of Quota 1 Applicants without Standby

	All applications (91,061 obs.)		Used sample (6,658 obs.)	
	Mean	Std. dev	Mean	Std. dev
A. Demographics				
Age	20.9	3.7	20.8	3.0
Male percent	43.9	–	40.1	–
Parental income (000's DKK)	302.4	157.4	325.1	167.8
Parental wealth (000's DKK)	361.8	731.2	416.5	798.2
Parental education (Years)	12.9	3.2	13.4	3.2
B. Application				
Grade	8.6	1.0	8.8	0.7
Nbr programs ranked	1.8	1.1	3.1	1.3
Accepted into reach	–	–	0.5	0.5
Rank of reach	–	–	1.2	0.5
Rank of safety	–	–	2.3	0.7
Rank of enrollment prog.	–	–	1.7	0.9
C. Reach-Safety pairs				
Grade distance from cutoff	–	–	0.0	0.3
Same field			0.5	–
Same institution	–	–	0.4	–
D. Earnings				
All	312.6	162.6	302.7	165.5
Women	279.4	141.7	282.7	150.5
Men	354.9	177.1	332.7	181.7
Enrollment cohort average	303.5	63.1	300.3	63.2
Enrollment cohort gender gap	47.8	36.1	39.8	29.8
Reach-Safety gap	–	–	13.0	60.7
Reach-Safety gap in gender gap	–	–	−1.3	38.6

Note: This table presents descriptive statistics for the population of students applying to enter university through the Quota 1 system of admission (without standby) for the first time in application years 1994 to 2006 (all applications; left) as well as the sample of 6,658 students who have a reach–safety programs pair in these application years (used sample; right). Panel A presents demographics. Parental income, wealth and education are the average values for both parents (or the values for a single parent). Panel B presents information typically found in a student's application. The student's score summarizes their high school grades. The number of priorities ranked refers to the number of programs for which the student applied. The maximum number of programs for which students can apply is eight. For the sample of students with a reach–safety programs pair, we present the rankings of the reach, safety and enrollment programs. Panel C describes the reach–safety program pairs. The distance from the cutoff represents the difference between the student's grade and the cutoff of the reach program; 60% of reach–safety program pairs are in the same field and 30% are at the same institution. Panel D presents information on earnings. The earnings are averaged 10 to 12 years after the application to university programs. We show the average earnings for all students and for the subsamples of women and men. The enrollment cohort earnings average is the average earnings of the cohort of students admitted to the same program–university major. The enrollment cohort earnings gender gap is the difference between the earnings of men and women for the cohort of students admitted to the same program–university major. The reach–safety earnings gap is the difference between the earnings of students admitted to their reach program and the earnings of students admitted to their safety program. The reach–safety earnings gender gap is the difference between the earnings gender gap for students admitted to their reach program and the earnings gender gap for students admitted to their safety program.

Table A2: Descriptive Statistics: Demographics of Quota 1 Applicants on Standby

	All applications (45,895 obs.)		Used sample (2,194 obs.)	
	Mean	Std. dev	Mean	Std. dev
A. Demographics				
Age	20.4	3.2	20.6	2.9
Male percent	39.1	–	36.0	–
Parental income (000's DKK)	321.6	163.6	358.6	179.6
Parental wealth (000's DKK)	385.4	755.6	478.6	820.9
Parental education (Years)	13.2	3.2	13.4	3.2
B. Application				
Grade	8.7	0.9	8.6	0.7
Nbr programs ranked	1.8	1.2	3.0	1.3
Accepted into reach	–	–	0.6	0.5
Rank of reach	–	–	1.1	0.4
Rank of safety	–	–	2.2	0.6
Rank of enrollment prog.	–	–	1.5	0.8
C. Reach-Safety pairs				
Grade distance from cutoff	–	–	0.0	0.3
Same field			0.5	–
Same institution	–	–	0.4	–
D. Earnings				
All	311.3	166.0	307.0	172.1
Women	288.7	147.8	291.5	160.4
Men	346.7	185.4	334.7	188.0
Enrollment cohort average	308.0	67.0	311.5	66.1
Enrollment cohort gender gap	42.6	34.0	40.2	32.4
Reach-Safety gap	–	–	24.8	63.4
Reach-Safety gap in gender gap	–	–	−3.5	46.2

Note: This table presents descriptive statistics for the population of students applying to enter university through the Quota 1 system of admission on standby for the first time in application years 1994 to 2006 (all applications; left) as well as the sample of 2,194 students who have a reach–safety programs pair in these application years (used sample; right). Panel A presents demographics. Parental income, wealth and education are the average values for both parents (or the values for a single parent). Panel B presents information typically found in a student's application. The student's score summarizes their high school grades. The number of priorities ranked refers to the number of programs for which the student applied. The maximum number of programs for which students can apply is eight. For the sample of students with a reach–safety programs pair, we present the rankings of the reach, safety and enrollment programs. Panel C describes the reach–safety program pairs. The distance from the cutoff represents the difference between the student's grade and the cutoff of the reach program; 60% of reach–safety program pairs are in the same field and 30% are at the same institution. Panel D presents information on earnings. The earnings are averaged 10 to 12 years after the application to university programs. We show the average earnings for all students and for the subsamples of women and men. The enrollment cohort earnings average is the average earnings of the cohort of students admitted to the same program–university major. The enrollment cohort earnings gender gap is the difference between the earnings of men and women for the cohort of students admitted to the same program–university major. The reach–safety earnings gap is the difference between the earnings of students admitted to their reach program and the earnings of students admitted to their safety program. The reach–safety earnings gender gap is the difference between the earnings gender gap for students admitted to their reach program and the earnings gender gap for students admitted to their safety program.

Table A3: Descriptive Statistics: Demographics of Quota 2 Applicants without Standby

	All applications (114,514 obs.)		Used sample (4,286 obs.)	
	Mean	Std. dev	Mean	Std. dev
A. Demographics				
Age	21.7	3.8	21.4	2.8
Male percent	34.9	–	37.2	–
Parental income (000's DKK)	284.1	143.0	334.1	168.6
Parental wealth (000's DKK)	320.2	680.4	458.4	840.8
Parental education (Years)	12.4	3.1	13.5	3.1
B. Application				
Grade	7.9	0.9	8.6	0.6
Nbr programs ranked	2.3	1.5	3.2	1.3
Accepted into reach	–	–	0.7	0.5
Rank of reach	–	–	1.2	0.6
Rank of safety	–	–	2.4	0.8
Rank of enrollment prog.	–	–	1.5	0.8
C. Reach-Safety pairs				
Grade distance from cutoff	–	–	−0.1	0.3
Same field			0.6	–
Same institution	–	–	0.3	–
D. Earnings				
All	290.7	145.4	316.8	161.8
Women	265.0	125.4	293.5	144.6
Men	338.5	166.4	356.0	180.7
Enrollment cohort average	284.3	47.5	304.6	62.3
Enrollment cohort gender gap	49.7	35.2	40.0	28.6
Reach-Safety gap	–	–	13.4	54.5
Reach-Safety gap in gender gap	–	–	−0.4	35.4

Note: This table presents descriptive statistics for the population of students applying to enter university through the Quota 2 system of admission (without standby) for the first time in application years 1994 to 2006 (all applications; left) as well as the sample of 4,286 students who have a reach–safety programs pair in these application years (used sample; right). Panel A presents demographics. Parental income, wealth and education are the average values of both parents (or the values for a single parent). Panel B presents information typically found in a student's application. The student's score summarizes their high school grades. The number of priorities ranked refers to the number of programs for which the student applied. The maximum number of programs for which students can apply is eight. For the sample of students with a reach–safety programs pair, we present the rankings of the reach, safety and enrollment programs. Panel C describes the reach–safety program pairs. The distance from the cutoff represents the difference between the student's grade and the cutoff of the reach program; 60% of reach–safety program pairs are in the same field and 30% are at the same institution. Panel D presents information on earnings. The earnings are averaged 10 to 12 years after the application to university programs. We show the average earnings for all students and for the subsamples of women and men. The enrollment cohort earnings average relates to the average earnings of the cohort of students admitted to the same program–university major. The enrollment cohort earnings gender gap is the difference between the earnings of men and women for the cohort of students admitted to the same program–university major. The reach–safety earnings gap is the difference between the earnings of students admitted to their reach program and the earnings of students admitted to their safety program. The reach–safety earnings gender gap is the difference between the earnings gender gap for students admitted to their reach program and the earnings gender gap for students admitted to their safety program.

Table A4: Descriptive Statistics: Demographics of Quota 2 Applicants on Standby

	All applications (68,838 obs.)		Used sample (6,060 obs.)	
	Mean	Std. dev	Mean	Std. dev
A. Demographics				
Age	22.0	3.7	21.0	2.6
Male percent	31.1	–	39.9	–
Parental income (000's DKK)	300.2	154.4	325.8	171.0
Parental wealth (000's DKK)	354.4	718.4	381.9	772.9
Parental education (Years)	12.8	3.2	13.5	3.2
B. Application				
Grade	8.0	0.9	8.6	0.7
Nbr programs ranked	2.6	1.6	3.1	1.4
Accepted into reach	–	–	0.6	0.5
Rank of reach	–	–	1.2	0.5
Rank of safety	–	–	2.3	0.7
Rank of enrollment prog.	–	–	1.5	0.9
C. Reach-Safety pairs				
Grade distance from cutoff	–	–	−0.1	0.3
Same field			0.5	–
Same institution	–	–	0.3	–
D. Earnings				
All	293.4	148.6	304.3	163.8
Women	273.4	130.9	284.0	146.5
Men	337.5	173.9	335.0	182.7
Enrollment cohort average	288.7	57.9	303.8	63.9
Enrollment cohort gender gap	41.9	28.2	40.9	27.0
Reach-Safety gap	–	–	18.0	60.6
Reach-Safety gap in gender gap	–	–	−0.4	34.6

Note: This table presents descriptive statistics for the population of students applying to enter university through the Quota 2 system of admission on standby for the first time in application years 1994 to 2006 (all applications; left) as well as the sample of 6,060 students who have a reach–safety programs pair in these application years (used sample; right). Panel A presents demographics. Parental income, wealth and education are the average values of both parents (or the values for a single parent). Panel B presents information typically found in a student's application. The student's score summarizes their high school grades. The number of priorities ranked refers to the number of programs for which the student applied. The maximum number of programs for which students can apply is eight. For the sample of students with a reach–safety programs pair, we present the rankings of the reach, safety and enrollment programs. Panel C describes the reach–safety program pairs. The distance from the cutoff represents the difference between the student's grade and the cutoff of the reach program; 60% of reach–safety program pairs are in the same field and 30% are at the same institution. Panel D presents information on earnings. The earnings are averaged 10 to 12 years after the application to university programs. We show the average earnings for all students and for the subsamples of women and men. The enrollment cohort earnings average relates to the average earnings of the cohort of students admitted to the same program–university major. The enrollment cohort earnings gender gap is the difference between the earnings of men and women for the cohort of students admitted to the same program–university major. The reach–safety earnings gap is the difference between the earnings of students admitted to their reach program and the earnings of students admitted to their safety program. The reach–safety earnings gender gap is the difference between the earnings gender gap for students admitted to their reach program and the earnings gender gap for students admitted to their safety program.

Table A5: Gender Earnings Gap First-stages

		Gender earnings gap		Separate earnings	
	All	Women	Men	Women	Men
$\mathbb{P}(p_i = r_i \| D_i) \times \overline{y_{r_i}}$	0.91	0.91	0.90		
	(0.02)	(0.03)	(0.03)		
$\mathbb{P}(p_i = s_i \| D_i) \times \overline{y_{s_i}}$	1.02	1.02	1.01		
	(0.03)	(0.04)	(0.04)		
$\mathbb{P}(p_i = r_i \| D_i) \times \overline{g_{r_i}}$		-0.23	-0.35		
		(0.05)	(0.06)		
$\mathbb{P}(p_i = s_i \| D_i) \times \overline{g_{s_i}}$		-0.28	-0.33		
		(0.05)	(0.07)		
$\mathbb{P}(p_i = r_i \| D_i) \times \overline{y_{r_i,m}}$				0.8	0.9
				(0.03)	(0.03)
$\mathbb{P}(p_i = s_i \| D_i) \times \overline{y_{s_i,m}}$				0.83	1.01
				(0.03)	(0.04)
$\mathbb{P}(p_i = r_i \| D_i) \times \overline{y_{r_i,w}}$				-0.82	-0.8
				(0.04)	(0.04)
$\mathbb{P}(p_i = s_i \| D_i) \times \overline{y_{s_i,w}}$				-0.83	-0.9
				(0.04)	(0.05)
R^2	0.79	0.79	0.80	0.93	0.95
N	10,944	6,676	4,268	6,676	4,268

Note: This table presents the results of the first-stage estimations of equation (2) presented in Table 4. We use the sample of 10,944 students who have a reach–safety programs pair in application years 1994 to 2006 and present the results for a bandwidth of +/– 0.5 around the admission cutoff of the reach program. The outcome variable for the regressions is the student's enrollment program's earnings averaged 10 to 12 years after admission. The first column uses the entire sample of applicants. The second and third columns separate the sample based on gender and use the average program earnings of enrollees and the earnings gap between males and females for the program of enrollment. The fourth and fifth columns separate the sample based on gender and use the average program earnings of men and women separately. The estimation controls for distance from cutoff fixed effects (bin FE), average earnings of the reach program, average earnings of the safety program, calendar year FE, and the probability of being accepted given last year's admission cutoff interacted with the average earnings of the reach and safety programs.

Table A6: Gender Earnings Gap Channels and Dynamics

	Specifications				
	(1)	(2)	(3)	(4)	(5)
A. 7 to 9 years after enrollment					
Program earnings (All)	0.44	0.44	0.44	0.43	0.44
	(0.13)	(0.13)	(0.12)	(0.12)	(0.12)
Within-program	-0.28	-0.28	-0.27	-0.31	-0.25
gender earnings gap	(0.16)	(0.16)	(0.16)	(0.16)	(0.15)
B. 10 to 12 years after enrollment					
Program earnings (All)	0.48	0.48	0.49	0.46	0.48
	(0.13)	(0.13)	(0.13)	(0.13)	(0.11)
Within-program	-0.45	-0.45	-0.42	-0.45	-0.35
gender earnings gap	(0.18)	(0.18)	(0.18)	(0.18)	(0.16)
C. 13 to 15 years after enrollment					
Program earnings (All)	0.54	0.53	0.53	0.52	0.51
	(0.17)	(0.17)	(0.17)	(0.17)	(0.16)
Within-program	-0.80	-0.81	-0.78	-0.8	-0.66
gender earnings gap	(0.23)	(0.23)	(0.23)	(0.22)	(0.21)
Married	NO	YES	YES	YES	YES
Nb. Of kids	NO	NO	YES	YES	YES
Graduated	NO	NO	NO	YES	YES
Any labor income	NO	NO	NO	NO	YES

Note: This table presents the results of estimating equation (2) by IVs for the gender earnings gap. We use the sample of 6,901 women who have a reach–safety programs pair in application years 1994 to 2006 and present the results for a bandwidth of +/– 0.5 around the admission cutoff of the reach program. The outcome variable for the regressions is the student's subsequent earnings averaged seven to nine years, 10 to 12 years and 13 to 15 years after admission in Panels A, B and C, respectively. The program earnings are averaged seven to nine years, 10 to 12 years and 13 to 15 years after the application to university programs in each of these panels. The first column repeats the baseline results presented in Table 6. The second column controls for a dummy variable indicating marital situation, the third column adds number of children fixed effects, the fourth column indicates whether the student has graduated from the reach program, and the fifth column adds a dummy variable for no labor income earnings.

Table A7: Gender Earnings Gap (All Admission Systems)

		Gender earnings gap		Gender-specific earnings	
	All	Women	Men	Women	Men
A. Instrumental variable estimation					
Program earnings (All)	0.27	0.48	-0.02		
	(0.08)	(0.09)	(0.15)		
Within-program gender earnings gap		-0.27	-0.01		
		(0.13)	(0.21)		
Program earnings (Women)				0.47	0.13
				(0.19)	(0.28)
Program earnings (Men)				-0.018	-0.018
				(0.16)	(0.24)
B. OLS benchmark					
Program earnings (All)	0.79	0.77	0.75		
	(0.02)	(0.02)	(0.04)		
Within-program gender earnings gap		-0.11	0.42		
		(0.05)	(0.08)		
Program earnings (Women)				0.68	-0.029
				(0.06)	(0.09)
Program earnings (Men)				0.15	0.84
				(0.05)	(0.07)
N	19,198	11,725	7,473	11,725	7,473

Note: This table presents the results of estimating equation (2) by IVs (Panel A) and by OLS (Panel B) for the gender earnings gap and for the subsamples of women and men. Each column presents the second stage of an IV regression. We use the sample of 19,198 students who have a reach–safety programs pair in application years 1994 to 2006 and present the results for a bandwidth of +/− 0.5 around the admission cutoff of the reach program. The outcome variable for the regressions is the student's subsequent earnings averaged 10 to 12 years after admission. The first column sets $\overline{g_{p_i}}$ to 0. The second and third columns separate the sample based on gender and use the average program earnings of enrollees and the earnings gap between men and women for the program of enrollment. The fourth and fifth columns separate the sample based on gender and use the average program earnings of men and women separately. The IV estimation controls for distance from cutoff fixed effects (bin FE), average earnings of the reach program, average earnings of the safety program, calendar year FE, the probability of being accepted given last year's admission cutoff interacted with the average earnings of the reach and safety programs. The OLS results control for parents' wealth, parents' income and parents' education, student grade, age and gender, and calendar year FE.

Table A8: Gender Earnings Gap (Only Quota 1, Including Standby)

		Gender earnings gap		Gender-specific earnings	
	All	Women	Men	Women	Men
A. Instrumental variable estimation					
Program earnings (All)	0.29	0.59	-0.12		
	(0.10)	(0.12)	(0.18)		
Within-program gender earnings gap		-0.24	0.12		
		(0.15)	(0.24)		
Program earnings (Women)				0.45	-0.13
				(0.24)	(0.33)
Program earnings (Men)				0.091	0.18
				(0.20)	(0.27)
B. OLS benchmark					
Program earnings (All)	0.75	0.78	0.67		
	(0.03)	(0.04)	(0.05)		
Within-program gender earnings gap		-0.17	0.36		
		(0.07)	(0.11)		
Program earnings (Women)				0.73	0.016
				(0.08)	(0.12)
Program earnings (Men)				0.093	0.73
				(0.07)	(0.10)
N	8,852	5,391	3,461	5,391	3,461

Note: This table presents the results of estimating equation (2) by IVs (Panel A) and by OLS (Panel B) for the gender earnings gap and for the subsamples of women and men. Each column presents the second stage of an IV regression. We use the sample of 8,852 students who have a reach–safety programs pair in application years 1994 to 2006 and present the results for a bandwidth of +/− 0.5 around the admission cutoff of the reach program. The outcome variable for the regressions is the student's subsequent earnings averaged 10 to 12 years after admission. The first column sets $\overline{g_{p_i}}$ to 0. The second and third columns separate the sample based on gender and use the average program earnings of enrollees and the earnings gap between men and women for the program of enrollment. The fourth and fifth columns separate the sample based on gender and use the average program earnings of men and women separately. The IV estimation controls for distance from cutoff fixed effects (bin FE), average earnings of the reach program, average earnings of the safety program, calendar year FE, the probability of being accepted given last year's admission cutoff interacted with the average earnings of the reach and safety programs. The OLS results control for parents' wealth, parents' income and parents' education, student grade, age and gender, and calendar year FE.

Table A9: Gender Earnings Gap Donut

		Gender earnings gap		Separate earnings	
	All	Women	Men	Women	Men
Instrumental variable estimation					
Program earnings (All)	0.26	0.42	0.11		
	(0.11)	(0.14)	(0.20)		
Within-program gender earnings gap		-0.41	-0.04		
		(0.18)	(0.25)		
Program earnings (Women)				0.67	0.32
				(0.27)	(0.36)
Program earnings (Men)				-0.28	-0.011
				(0.23)	(0.29)

Note: This table presents the results of estimating equation (2) by IVs in a donut-hole analysis that excludes applicants with a score exactly on the admission cutoff for their reach program. Each column presents the second stage of an IV regression. We use the sample of women who have a reach–safety programs pair in application years 1994 to 2006 and present the results for a bandwidth of +/− 0.5 around the admission cutoff of the reach program. The outcome variable for the regressions is the student's subsequent earnings averaged 10 to 12 years after admission.