THE RETURNS TO STEM PROGRAMS FOR LESS-PREPARED STUDENTS

KEVIN NG*
EVAN RIEHL+

SEPTEMBER 2020

Abstract. We examine how returns to enrolling in science, technology, engineering, and mathematics (STEM) programs vary with students' academic preparation. We match data on STEM admissions at a Colombian flagship university to nationwide college and earnings records. Our identification strategy combines a regression discontinuity design with variation in admission quotas. We find that less-prepared students were less likely to complete a STEM degree than their more able peers, but they had *larger* earnings returns to enrolling. Our results suggest that policies that encourage less-prepared students to enroll in STEM programs can yield large but unevenly distributed earnings gains.

*Ng: Department of Economics, Cornell University, 404 Uris Hall, Ithaca, NY 14853 (email: kcn38@cornell.edu). +Riehl (corresponding author): Department of Economics, Cornell University, 266 Ives Hall, Ithaca, NY 14853 (email: eriehl@cornell.edu). For helpful comments we thank Natalie Bau, David Deming, Nicolás de Roux, Jorge Garcia, Matthew S. Johnson, Michael Lovenheim, Eleonora Patacchini, Seth Sanders, and Christopher Walters. We are grateful to Universidad del Valle for assistance with the data. All errors are our own.

Many countries have set goals to increase the number of college students who earn degrees in science, technology, engineering, and mathematics (STEM). These policies are partly motivated by the fact that STEM majors often have higher earnings than graduates from other fields (Altonji et al., 2012), but these returns are typically calculated using individuals who completed a degree. STEM programs also have high dropout rates, especially among students with relatively less academic preparation (Arcidiacono, 2004; Stinebrickner and Stinebrickner, 2014). If students who are drawn into STEM programs by policy initiatives are less prepared on average, it is unclear whether they would also have large returns.

This paper asks how the returns to enrolling in selective STEM programs vary with students' academic preparation. We exploit data on admissions, graduation, and earnings for applicants to selective STEM programs at a flagship university in Colombia. Our identification strategy combines a regression discontinuity design based on the flagship's admission thresholds with variation in admission quotas across cohorts. This allows us to estimate STEM graduation rates and earnings returns for students who differ by more than ten percentile points in the national distribution of academic ability for high school graduates.

Our research question is inspired by two types of policies that aim to increase the number of college graduates in STEM fields. First, many countries have initiatives to induce more students to pursue STEM majors through scholarships or expanded program capacity (e.g., Holdren, 2013). Since STEM programs are among the most challenging majors at many colleges, the students who are affected by these policies tend to be relatively less prepared for STEM coursework. Second, there is an ongoing debate on how affirmative action in university admissions affects STEM degree attainment. For example, Arcidiacono et al. (2016) present evidence that the number of STEM graduates would increase—particularly in underrepresented minority groups—if students were better "matched" to colleges based on ability. To evaluate these policies, it is important to understand how both graduation rates and earnings returns vary with individuals' academic preparation.

Our paper proceeds as follows. Section 1 motivates our analysis by presenting descriptive facts on the relationship between ability, field of study, degree attainment, and earnings. We use data from Colombia's national standardized ICFES exam to define a measure of ability for nearly all high school graduates in the country from 1998–2004. We match these data to enrollment/graduation records from a higher education census, and to administrative earnings records for the year 2017.

We show that lower-ability students were less likely to enroll in and complete STEM programs than their more able peers, but they had especially large earnings premiums to STEM degrees. Only one in five college enrollees with below-median ability chose a STEM program, and their mean graduation rate in those programs was 18 percent. But less able students who completed a STEM degree earned 10 percent more on average than individuals with similar ability in the next-highest paying field. The STEM earnings premium decreases with ability, and at the highest ability levels business graduates earned more than STEM graduates. Further, the relationship between earnings and ability is flatter among graduates from the same STEM programs than in other higher-paying fields. These patterns suggest that STEM degrees may be particularly beneficial to less-prepared students, but our descriptive analysis cannot convincingly separate causal effects from selection into STEM completion.

In Section 2, we turn our attention to one university where we can estimate the causal returns to STEM programs. This university is a public flagship institution in Cali, Colombia called *Universidad del Valle*. The flagship offers roughly 60 degree programs each year including STEM programs in engineering and natural sciences. We matched our administrative datasets to admission records for applicants to each flagship program in 1999–2004. This allows us to observe applicants' enrollment choices, graduation outcomes, and formal sector earnings 13–17 years later.

Section 3 develops a framework that shows how returns to the flagship STEM programs can vary with ability. Our estimates pertain to a group of "compliers" who would enroll in a flagship STEM program if they are admitted, and who vary in the program they choose when rejected. All else equal, earnings returns to STEM enrollment would increase with ability if more able students are more likely to graduate. But there are several reasons why less-prepared students could have larger returns even if they have lower graduation rates. Less-prepared students may enroll in programs with lower earnings potential if they are not admitted to the flagship. Less-prepared students could also have larger value added from a completed STEM degree if they learn more along the way.

Section 4 uses a regression discontinuity (RD) design to estimate average returns to enrolling in the flagship's STEM programs. Students apply to specific programs at the flagship, and admission is solely determined by scores on the ICFES exam. We develop an RD design that estimates the returns to enrolling in STEM programs for students on the margin of admission. We find that the flagship's STEM programs are unique in that they have the lowest graduation rates but the highest mean earnings returns to enrollment. Across all engineering and natural science programs, the mean graduation rate for marginal admits was only 34 percent, but enrolling in these programs increased average monthly earnings in 2017 by 14 percent. In other flagship programs, the graduation rate for marginal admits was 50 percent, and the mean earnings return was close to zero.

In Section 5, we use two different empirical strategies to examine how STEM returns vary with academic preparation. First, we estimate heterogeneity in the RD coefficients by exploiting the fact that our data contain multiple measures of pre-college ability. Flagship admissions are based on a weighted average of up to nine subject scores on the ICFES exam. Thus marginal admits vary in ability as defined by outcomes other than the admission score. Our benchmark measure of ability averages the ICFES subject scores based on their predictive power for college graduation, which is similar to Cunha et al. (2010)'s anchoring approach. This allows us to estimate STEM returns for applicants who vary by roughly 15 percentile points in the national distribution of graduation-anchored ability.

Our second empirical strategy uses variation in the size of the flagship's admission quotas across cohorts. During our sample period, the flagship usually admitted about 60 applicants to each program. In a few STEM programs, however, changes in admission policies caused these quotas to expand to over 120 applicants in certain cohorts (de Roux and Riehl, 2019b). We develop an RD differences-in-differences design that exploits large changes in quotas across programs and cohorts, and show that these changes caused mean ability near the threshold to vary by nearly ten percentile points.

Our main result is that less-prepared students were less likely to graduate from the flagship's STEM programs than their more able peers, but they had *larger* earnings returns to enrolling. This pattern appears in both of our empirical strategies. A ten percentile point decrease in ability is associated with a ten percentage point decline in the flagship graduation rate, but a roughly 20 percent increase in the mean earnings return. This result is unique to the flagship's STEM programs, as graduation and earnings outcomes varied less systematically with ability in other degree programs.

Lastly, in Section 6 we explore mechanisms for the inverse relationship between graduation rates and earnings returns across ability levels. There is little evidence that students who dropped out of the STEM programs had large returns to enrolling. Instead, we find suggestive evidence for two mechanisms. First, lower-ability students were less likely to enroll in other STEM programs when they were rejected, so flagship admission had a similar effect on the probability that high- and lower-ability students attained *any* STEM degree. Second, earnings and ability are only weakly related in the population of flagship graduates, which suggests that the value added of a STEM degree was larger for less-prepared students.

Our findings show that policies that encourage students to enroll in STEM programs can yield large earnings returns, even if the affected students are less prepared for STEM coursework. These policies are motivated by a perceived shortage of STEM workers (Carnevale et al., 2011; Deming and Noray, 2019) and by evidence that STEM skills are important for growth and innovation (Peri et al., 2015; Bianchi and Giorcelli, 2019). Our results suggest that students affected by these policies may have even larger mean earnings returns than existing STEM enrollees, but these gains can be unevenly distributed given low degree attainment rates. An important caveat is that our estimates do not capture general equilibrium effects that might arise from very large increases in STEM enrollment (Bianchi, 2020).

Our paper also informs debates on how admission policies like affirmative action affect the production of STEM skills. Like many other papers, we find that academic preparation is a strong predictor of whether students complete a STEM program (Sabot and Wakeman-Linn, 1991; Ost, 2010; Arcidiacono et al., 2012; Stinebrickner and Stinebrickner, 2014; Arcidiacono et al., 2016). However, we also find that less-prepared students were less likely to enroll in other STEM programs when they were rejected by the flagship. Thus raising admission standards by eliminating affirmative action may not increase the likelihood that lower-ability students attain STEM degrees. Many of these papers are also unable to examine labor market outcomes. Our results show that higher admission standards can reduce the mean earnings return for STEM enrollees. They also illustrate why less-prepared students may choose to enroll in selective STEM programs even if the likelihood of graduation is low.

Our earnings results corroborate work on the returns to field of study (Altonji et al., 2012, 2016), although our analysis differs from this literature in several ways. Our estimates do not isolate the returns of one field relative to another (Kirkebøen et al., 2016) because rejected applicants in our sample chose a heterogeneous set of programs, and because many did not earn a degree. Nonetheless, our exploration of mechanisms suggests that ability variation in STEM returns is driven by differences in the value added of a completed degree. This contributes to work on how STEM returns vary with ability (Altonji, 1993; Arcidiacono, 2004; Webber, 2014; Kinsler and Pavan, 2015), which has yielded mixed evidence.

Lastly, our paper is related to research that uses RD designs to estimate returns to selective colleges and majors (Hoekstra, 2009; Saavedra, 2009; Hastings et al., 2013; Zimmerman, 2014, 2019; Kirkebøen et al., 2016; Goodman et al., 2017; Canaan and Mouganie, 2018; Anelli, 2018; Smith et al., 2020). We contribute to this work by exploiting additional variation to estimate returns at different points in the distribution of ability.[1] Our paper also differs from many in this literature because Colombia has a decentralized college admission system, and our data allow us to observe enrollment in nearly all programs in the country. Since students' application behavior can vary widely in decentralized systems (Hoxby and Avery, 2013; Pallais, 2015; Machado and Szerman, 2017), it is important to understand how marginal STEM enrollees vary in their counterfactual program choices.

1. Motivation

This section discusses policies for increasing the number of STEM graduates that motivate our analysis of ability variation in STEM returns.[2] We then use Colombian administrative data to describe how enrollment, graduation, and earnings outcomes vary by ability and

[1] Our RD heterogeneity analysis is broadly similar to Angrist and Rokkanen (2015)'s method for identifying RD effects away from the threshold under a conditional independence assumption. Our identification strategy based on admission quotas relies on additional policy variation.

[2] Throughout the paper, we use the terms "ability" and "academic preparation" interchangeably to refer to individuals' accumulated human capital at the end of high school.

degree field. We show that STEM enrollment and graduation rates were lower for less-prepared students, but they had the largest earnings premiums to a completed STEM degree.

1.1. **Policies to increase the number of STEM graduates.** Our analysis is motivated by two types of policies that aim to increase the number of college graduates with STEM degrees. First, many countries have devoted resources to expand university STEM programs. These resources fund STEM scholarships, faculty hiring, and facilities (e.g., Holdren, 2013). Since STEM programs are among the most challenging programs at selective colleges, such efforts are likely to attract students from other majors or institutions who are lower in the distribution of academic preparation.[3]

Second, Arcidiacono et al. (2016) argue that changes in admission standards at selective colleges may increase the number of students who earn STEM degrees—particularly among underrepresented minority groups. The authors show that students are more likely to switch out of STEM programs when their own academic preparation is significantly below that of their classmates. Their analysis suggests that the number of STEM graduates would increase if students were better "matched" to universities based on pre-college ability. This relates to debates on affirmative action in college admissions, which leads to large gaps in academic preparation between racial/ethnic groups at selective colleges.

The desirability of both of these policies depends on whether relatively less-prepared students have significant earnings returns to STEM programs. Several papers have shown that lower-ability students are less likely to persist in STEM programs conditional on initial interest (e.g., Arcidiacono, 2004; Stinebrickner and Stinebrickner, 2014). Research on the returns to field of study often finds large earnings premiums to a completed STEM degree (Altonji et al., 2012; Kirkebøen et al., 2016), but there is less evidence on how these returns vary with ability. Our goal in this paper is to jointly analyze the relationship between ability, graduation, and earnings for students who enroll in STEM programs.

1.2. **Colombian administrative data and institutional background.** We begin with a descriptive analysis of heterogeneity in college and labor market outcomes by ability and degree field. For this we use three administrative datasets from the country of Colombia.

Our first dataset is from a national standardized exam called the ICFES, which all Colombian students are required to take to apply to college. The ICFES is similar to the SAT exam in the U.S., but it is taken by nearly all high school graduates in the country. It also contains more detailed subjects; over the past two decades the ICFES has included 7–10

[3] Policies that increase college resources seek to expand STEM enrollment given a fixed distribution of ability. Our paper does not directly address other policies that help students develop STEM skills at younger ages.

different subject tests.[4] Our dataset includes subject scores and demographic characteristics for all students who took the exam in 1998–2004.

Second, we use Ministry of Education data on enrollment and graduation at nearly all colleges in the country. Colombia has a wide range of public and private colleges with varying selectivity and degree offerings. Most of its 33 regions have a public flagship university similar to those in the U.S. Flagships are typically the most selective institutions in each region and are much less expensive to attend than comparable private colleges. Like in the U.S., Colombia has a decentralized system of college admissions; students apply separately to each institution, and colleges set their admission criteria. A difference is that Colombian students apply to institution/major pairs that we call "programs." Our data include all students who enrolled in college programs tracked by the Ministry in 1998–2012. We observe each student's institution, field of study, dates of entry and exit, and graduation outcome.

Finally, we use data on earnings from tax records managed by the Ministry of Social Protection. These data provide monthly earnings in 2017 for any individual who worked in the formal sector. Our data do not include earnings from informal firms that are not registered with the Ministry. The informal sector is a substantial portion of the Colombian labor market, although it is less important for college-educated workers.[5] In our analysis we examine effects on both formal employment—defined as appearing in the Ministry's data—and on earnings. Formal employment is likely to lead to higher earnings, as mean hourly wages are roughly 50 percent higher in the formal sector (de Roux and Riehl, 2019a).

We link the administrative datasets using individuals' names, birthdates, and ID numbers. Appendix B.2 provides details on the data coverage and merge process. The combined dataset includes the college outcomes of nearly all Colombian high school graduates in 1998–2004, and their earnings measured 13–19 years later.

1.3. **Enrollment and graduation patterns.** To explore variation in outcomes by ability, we use ICFES scores to define a measure of pre-college academic preparation that we use throughout the paper. For this we regress an indicator for college graduation on ICFES subject scores in a sample of all college enrollees in our data. Our ability measure is the predicted values from this regression, which we compute for all exam takers. This measure averages ICFES subject scores based on their predictive power for college graduation, which is similar to Cunha et al. (2010)'s approach of anchoring test scores to an outcome of interest.[6] We convert this measure to percentile units within the population of all exam takers in each

[4] Appendix Table A2 lists the subjects during the period of our data. The ICFES is now named Saber 11.

[5] In our sample, formal sector employment rates are 81 percent for college graduates and 63 percent for non-graduates. The remaining individuals are either informally employed or out of the labor force.

[6] Our main results are similar if we define ability as the simple average of ICFES subjects scores. We use a graduation-anchored measure to focus directly on how earnings outcomes vary for students with different likelihoods of completing a degree. Appendix Table A2 provides details on this anchoring regression.

year. Thus our measure represents an individual's position in the distribution of graduation-anchored ability for nearly all high school graduates in the country.

Figure 1 shows how college and earnings outcomes varied with ability. The x-axis in each panel is our measure of ability in percentile units from zero to one. Lines depict non-parametric relationships between ability and the outcome listed in each panel title.

Panel A shows that there is a strong relationship between ability and college enrollment. We focus on enrollment in bachelor's degree programs, which typically have on-time durations of 4–5 years.[7] In the 1998–2004 high school graduation cohorts, nearly 80 percent of individuals at the top of the ability distribution enrolled in a bachelor's program by 2012. Bachelor's program enrollment rates were roughly 10 percent at the lowest ability levels.

Higher-ability students were also more likely to enroll in STEM programs. Panel B shows the distribution of fields conditional on college enrollment using the Ministry of Education's categorization of programs. The solid line includes STEM programs, which include those in the engineering and natural science areas. The dashed lines represent programs in business, law/social sciences, health, and education/arts.[8] The fraction of enrollees who chose a STEM program rises from 20 percent at the bottom of the ability distribution to nearly 60 percent at the top. Enrollment in each of the other four degree categories is mostly decreasing in ability. Less-prepared students were roughly equally likely to choose each of five fields.

STEM programs typically had the lowest graduation rates, and degree completion was especially rare for less-prepared STEM enrollees. Panel C shows the fraction of enrollees who completed their program by 2012 in each degree field. At the highest ability levels, 64 percent of students who enrolled in a STEM program earned a degree. Graduation rates were 25 percent or less for STEM enrollees with below-median ability. STEM programs had the lowest completion rates at most ability levels, and the gap in graduation rates between STEM and other fields is generally decreasing in ability. In the bottom ability quartile, the STEM graduation rate is 10 percentage points lower than the mean completion rate in other programs. This gap is only 6 percentage points in the top quartile of ability.

The enrollment and graduation patterns in Figure 1 are similar to those in many other countries. In the University of California system, for example, Arcidiacono et al. (2016) show that students who intend to major in the sciences have higher SAT scores and high school GPAs than students with other initial majors. Across the UC campuses, science degree completion rates range from 40–60 percent for students in the top quartile of academic preparation, and are 10–20 percent for students in the bottom quartile.

[7] Many colleges also offer 2–3 year technical programs, but we do not focus on these in our analysis.

[8] The Ministry's social sciences category includes some humanities programs, but the dominant majors are law, psychology, communications, and sociology. We include agronomy programs in the education/arts category because they have similar graduation rates and mean earnings.

1.4. **Earnings patterns.** Colombian students who completed a STEM degree had high average earnings, and the STEM premium relative to other degrees is larger for less-prepared individuals. Panel D of Figure 1 shows the relationship between ability and log monthly earnings in 2017 for individuals who attained a college degree. At most ability levels, STEM graduates had the highest mean earnings, and the earnings premium over the next highest-paying fields is roughly ten percent. The STEM earnings premium narrows above the 80$^{\text{th}}$ percentile of ability, and mean earnings are higher for business graduates than for STEM graduates at the highest ability levels.

Table 1 shows how earnings vary with ability among graduates from the same programs. Column (A) shows mean monthly earnings converted to 2017 U.S. dollars in a sample of all 1998–2004 high school graduates who completed a bachelor's degree. Column (B) reports the coefficient from a regression of log earnings on an individuals' ability percentile *within* their graduation cohort. Specifically, we renormalize our graduation-anchored ability measure to reflect an individual's percentile in the distribution of students who graduated from the same college program in the same year. The coefficients in column (B) can therefore be interpreted as the mean difference in log earnings between graduates from each program with the highest and lowest pre-college ability. The first row of Table 1 displays pooled estimates across all degrees. The other rows show separate estimates for our five degree fields.

The results in Table 1 show that STEM programs have high mean earnings and relatively flat earnings/ability gradients. In 2017, college graduates had average monthly earnings equivalent to 872 U.S. dollars, and the most able graduates from each program earned 20 percent more (0.18 log points) than the least-prepared graduates. In STEM programs, the earnings gap between the most- and least-prepared students is only 15 percent despite higher mean earnings. Earnings and ability are more strongly related in business and law/social science programs. Health and education/arts programs have earnings/ability gradients similar to that in STEM, although graduates in these fields have significantly lower mean earnings.

Columns (C)–(D) show that the relationship between earnings and ability is even weaker for STEM graduates from flagship universities. These columns are similar to columns (A)–(B), but we restrict the sample to programs at 24 public flagships across the country. Flagship graduates had slightly lower average earnings because flagships are geographically dispersed, while other colleges are disproportionately located in the capital city of Bogotá. But the earnings/ability gradient for flagship students is about one-third less than that for the full population of college graduates. This is driven by a flatter gradient in flagship STEM programs, where the most able graduates earned only seven percent more than the least-prepared graduates in 2017. Earnings and ability are also less related in other flagship programs relative to those at other colleges, but the magnitude of this difference is smaller.

Figure 1 and Table 1 suggest that STEM degrees may be especially beneficial for students with relatively less academic preparation. This pattern is most pronounced at flagship universities, where the least-prepared STEM graduates have earnings outcomes that are not far behind those of their peers. This result is broadly consistent with Chetty et al. (2017)'s finding that family background and earnings outcomes are weakly related among students who attended the same U.S. colleges. Our findings suggest that STEM degrees may play an important role in flattening this relationship.

These findings help to justify policies that induce relatively less-prepared students to enroll in selective STEM programs, but our descriptive analysis is not conclusive on this matter. Since academic preparation is particularly important for graduation in STEM programs, it is important to understand the net effects of low completion rates and high earnings premiums. Further, Figure 1 and Table 1 do not necessarily reflect the causal effects of STEM degrees. For example, lower-ability students who enroll in and complete STEM program are likely to be positively selected on unobservable dimensions of ability. These caveats motivate the rest of our paper, which analyzes a setting in which we can more credibly identify ability variation in the returns to *enrolling* in STEM programs.

2. Flagship data and sample

This section provides background on the flagship university in our analysis and its admission procedures. We also describe our data, sample, and classification of STEM programs.

2.1. **The flagship university.** The rest of our paper focuses on a public flagship university in Colombia called *Universidad del Valle*, or *Univalle* for short. Univalle is located in Cali—the country's third largest city and the capital of the Valle del Cauca region. Like other flagships, Univalle is much less expensive than comparable private colleges and offers tuition discounts to low-income students. It is the largest and most selective university in its region; the median Univalle enrollee in our sample scored at the 92$^{\text{nd}}$ percentile on the national standardized ICFES exam, and its yield among admitted students was 90 percent. Univalle is slightly larger than the average Colombian flagship, and is similar in mean selectivity.

Univalle offers roughly 60 undergraduate majors each year in STEM and other fields. Students apply to specific programs, and admission is based solely on the ICFES exam. Univalle's admission scores are weighted averages of applicants' scores on the ICFES subject tests, with weights that vary across programs; for example, STEM programs place more weight on math and natural science subjects. Applicants with the highest admission scores are admitted up to a cutoff that is determined by the number of available slots in each program. During the period of our data, Univalle typically admitted 45–65 students per cohort in each program. Section 5 describes several admission policies that caused these quotas to increase, which we use to identify heterogeneity in returns to the flagship programs.

2.2. **Sample.** We collected data on all applicants to Univalle's undergraduate programs from Fall 1999 to Spring 2004. These data include applicants' admission scores and admission decisions. We linked the flagship records to the national administrative datasets described in Section 1.2 (see Appendix B.2 for details). The resulting dataset allows us to observe applicants' admission outcomes, enrollment choices at nearly all Colombian colleges, graduation outcomes, and earnings measured 13–17 years later.

We define STEM programs to be the flagship's engineering and natural science majors. Univalle groups its programs into different faculty areas as shown in column (A) of Table 2. Column (B) provides examples of programs in each area. We classify programs in the engineering and natural science areas as STEM. This includes Biology, Chemistry, Mathematics, Physics, Statistics, and a dozen different engineering programs. We define all other programs as non-STEM, including Economics and some health programs where the classification is less clear. Our definition follows the standard classification of STEM majors in Canada, which has pre-professional undergraduate health programs similar to those in Colombia.[9] Below we also show that programs like Medicine and Dentistry have much higher graduation rates than similarly selective programs in engineering and natural sciences.

Our sample includes all bachelor's degree programs where we can identify the effects of flagship admission. Our initial dataset includes 74 different degree programs over the 1999–2004 academic years. We exclude 2–3 year programs that terminate in technical degrees and programs that did not use ICFES scores for admissions (e.g., Music). We also drop program/cohort pairs for which all applicants were accepted. Our final sample includes 48 different bachelor's programs, as shown in column (C) of Table 2. This includes 242 unique program/cohort pairs (column (D)). Appendix Table A1 lists the full set of programs in our sample, and Appendix B.3 provides details on how we construct the sample.

Our flagship sample includes all applicants who faced the standard admission criterion in each pool. We exclude applicants in special admission groups who were not subject to the primary admission thresholds (e.g., disabled or indigenous applicants), as well as applicants with missing ICFES scores. This follows Abdulkadiroğlu et al. (2014)'s approach in defining "sharp" RD samples, in which admission is equivalent to having a score above the threshold. Column (E) shows that our final sample includes roughly 40,000 applicants. Approximately 12,500 of these applicants were admitted (column (F)), for an overall admission rate of 31 percent. 90 percent of admitted students accepted their admission offer (column (G)).

[9] Statistics Canada classifies programs as either STEM or BHASE—business, health, humanities, arts, social science, and education (Statistics Canada, 2017). Economics is included in the BHASE group.

Column (H) in Table 2 shows that the mean flagship admit scored at the 87$^{\text{th}}$ percentile of the ICFES exam. This column uses our graduation-anchored measure of ability, as defined in Section 1.3. STEM program admits had higher average ability than those in other programs. Below we show that there is considerable variation in ability both across STEM programs and within each application pool.

3. Framework

This section presents a framework that describes the returns to the flagship's STEM programs that are relevant for our analysis and shows how they can vary with ability.

3.1. **Mean returns to flagship STEM enrollment.** We consider a population of high school graduates indexed by i, and each graduate has pre-college ability that we denote by α_i.[10] Individuals can choose from a large number of college programs $p \in P$, where programs are defined by both an institution and a field of study. The set P also includes the option of not enrolling in college at all. We allow for the possibility that individuals add to their skill merely by *enrolling* in a college program. Specifically, we let v_{ip}^e denote individual i's potential skill value added from enrolling in program p. This term reflects factors like learning in first-year courses or training opportunities that are enabled by college enrollment.

Students can gain additional skill if they complete their college program. We let g_{ip} denote individual i's potential graduation outcome in program p. In other words, $g_{ip} = 1$ for individuals who would successfully complete the program if they enrolled, and $g_{ip} = 0$ for those who would drop out. We let v_{ip}^g represent the additional skill value added that individual i would get if they graduate from program p.[11]

Importantly, each of these three potential outcomes—v_{ip}^e, g_{ip}, and v_{ip}^g—can depend on an individual's ability, α_i. For simplicity, we do not explicitly model the possibility that these outcomes depend on peer composition in an individual's program. Below we discuss how such peer effects would influence our empirical results.

After college, individuals enter a competitive labor market and earn a wage equal to their skill. We abstract from additional wage growth from on-the-job training. Under these assumptions, individual i's potential log wage from enrolling in program p is given by:

$$w_{ip} = \alpha_i + v_{ip}^e + g_{ip} v_{ip}^g. \tag{1}$$

An individual's wage is equal to $\alpha_i + v_{ip}^e + v_{ip}^g$ if they complete program p, and it is equal to $\alpha_i + v_{ip}^e$ if they drop out of the program.

[10] We assume ability is unidimensional for simplicity and to align with our empirical analysis. A more complex but realistic framework would allow individuals to vary in ability for different fields (Altonji, 1993).
[11] We assume v_{ip}^e, $v_{ip}^g \geq 0$ for all p, and that there is no value added for the option of not attending college.

Our analysis is relevant for a population of *compliers* for a flagship STEM program that we denote by s. Compliers are students who would enroll in program s if and only if they are offered admission. If they are not admitted, compliers enroll in their next-choice program that we denote by $c(i) \in P$. Next-choice programs can vary across individuals in the complier group, and they may differ from program s in institution and/or field of study.

In Section 4, we begin our empirical analysis by estimating mean wage returns to enrolling in the flagship's STEM programs. In notation, this return is $E[w_{is} - w_{i,c(i)}]$, where the expectation is defined over all compliers who are close to the flagship's admission threshold. Using equation (1), one can decompose the mean return to program s into three terms:

(a) The average skill gain to *enrolling* in program s relative to next choice programs, $E[v^e_{is} - v^e_{i,c(i)}]$;
(b) The graduation rate in program s relative to next-choice programs, $E[g_{is} - g_{i,c(i)}]$;
(c) The average skill gain to *graduating* from program s relative to next choice programs, $E[v^g_{is} - v^g_{i,c(i)}]$.

Appendix B.1 shows the full expression for this return as a function of these terms.

Our parameter of interest differs in two key ways from that in research on returns to field of study (Altonji et al., 2012). First, program s compliers vary in their next-choice programs, $c(i)$, whereas field of study research focuses on the returns to one major relative to another. Second, our estimand includes earnings outcomes for individuals who do not graduate. Heterogeneity in enrollment choices and graduation rates is important for understanding ability variation in STEM returns.

3.2. **Heterogeneity in STEM returns by ability.** Section 5 explores how returns to the flagship's STEM programs vary with individual ability. In notation, this estimand is $dE[w_{is} - w_{i,c(i)}|\alpha_i = \alpha]/d\alpha$—the change in the mean return to program s from an increase in ability, α.

All else equal, wage returns to the flagship's STEM programs would increase with ability if more able students have higher graduation rates. This effect alone would lead wage returns to increase with ability if there is a large value added to a flagship STEM degree.

But STEM returns are not necessarily larger for more able students because other potential outcomes can also vary with ability. In our framework, there are three reasons why less-prepared students could have larger returns to program s despite lower graduation rates:

(a) There is a skill return to *enrolling* in program s that decreases with ability,

$$\frac{dE[v^e_{is} - v^e_{i,c(i)}|\alpha_i = \alpha]}{d\alpha} < 0;$$

(b) Lower-ability students choose counterfactual programs with less degree value added,

$$\frac{dE[v^g_{i,c(i)}|\alpha_i = \alpha, g_{i,c(i)} = 1]}{d\alpha} > 0;$$

(c) Lower-ability students have greater value added to a degree from program s,

$$\frac{dE[v^g_{is}|\alpha_i = \alpha, g_{is} = 1]}{d\alpha} < 0.$$

Appendix B.1 provides a formal statement and proof of this claim using the expression for $dE[w_{is} - w_{i,c(i)}|\alpha_i = \alpha]/d\alpha$.

We explore the importance of these three potential mechanisms in our empirical analysis. Case (a) could arise if enrolling in flagship STEM programs provides access to training opportunities with especially large benefits for lower-ability students. Case (b) could arise if less able students are more likely to choose degree programs with low earnings potential when they are rejected from the flagship. Case (c) reflects the possibility that flagship STEM degrees help less-prepared students "catch up" to their more able peers if, for example, they learn more in the process of obtaining a degree.

An important consideration is that a student's classmates may also influence their potential outcomes, and class composition may change if STEM policies are large enough in scale. For example, the propensity to graduate or the value added of a degree could depend not just on an individual's own ability, but also on the abilities of their classmates. Such effects could arise from peer interactions (Sacerdote, 2001) or from professor responses to classroom composition (Duflo et al., 2011). These mechanisms would lead large-scale policies to have different effects from policies that induce smaller changes in STEM enrollment (Bianchi, 2020). Below we discuss the potential role of peer effects and the degree to which they are reflected in our empirical estimates.

4. Mean returns to flagship programs

This section presents average graduation rates and earnings returns for marginal admits to the flagship's programs. We first describe our regression discontinuity design and identification assumptions. We then show that STEM programs had the lowest graduation rates but the highest earnings returns among flagship programs.

4.1. **Regression discontinuity specifications.** We use reduced-form and two-stage least squares (2SLS) models to estimate the effects of admission and enrollment in flagship programs. Our reduced-form model is a stacked regression discontinuity (RD) regression:

$$Y_{ip} = \theta D_{ip} + \alpha x_{ip} + \psi D_{ip} x_{ip} + \gamma_p + \epsilon_{ip} \quad \text{if } |x_{ip}| \leq h. \tag{2}$$

The dependent variable is an outcome for individual i who applied to the flagship in application pool p. Application pools are defined by a program and semester of application, and in some years, by the applicant's version of the ICFES exam.[12] The variable of interest, D_{ip}, is an indicator for having an admission score above the threshold for that application pool. We include a linear spline in the running variable, x_{ip}, which is an individual's rank in their application pool normalized to equal zero for the last student above the threshold. We also include fixed effects for each application pool, γ_p, as is standard in stacked RD specifications (Pop-Eleches and Urquiola, 2013; Abdulkadiroğlu et al., 2014).

Our regressions include the subset of applicants whose admission ranks, x_{ip}, are within h positions of the admission thresholds. Our benchmark model uses $h = 30$, which is approximately the mean of the Calonico et al. (2014) bandwidths across all dependent variables. Appendix Tables A3 and A4 show that our main results are similar using bandwidths of 15 or 45 positions, and also if we use the Calonico et al. (2014) bandwidth for each outcome.[13] We cluster standard errors at the individual level; the level of treatment is an individual/application pool pair, but clustering addresses the fact that some individuals appear in our sample multiple times because they reapplied to the flagship.

In the reduced-form specification (2), our coefficient of interest, θ, measures the effect of *admission* to the flagship program. We use this model to examine enrollment rates for marginal admits to the flagship's programs, as well as effects of admission on graduation and labor market outcomes.

To examine the returns to *enrolling* in a flagship program, we use a 2SLS specification. The second-stage regression in this model is:

$$Y_{ip} = \beta E_{ip} + \tilde{\alpha} x_{ip} + \tilde{\psi} D_{ip} x_{ip} + \tilde{\gamma}_p + \tilde{\epsilon}_{ip} \quad \text{if } |x_{ip}| \leq h. \tag{3}$$

This specification is similar to equation (2), except the variable of interest is an indicator equal to one if the applicant enrolled in the flagship program they applied to, which we denote by E_{ip}. We use the above-threshold indicator, D_{ip}, to instrument for E_{ip}.[14]

In 2SLS regressions with log earnings as the dependent variable, Y_{ip}, our coefficient of interest, β, identifies the return to enrolling in a flagship program. Under the assumptions discussed below, this coefficient can be interpreted as the average return for a population of flagship "compliers," which is the subset of marginal admits whose enrollment was affected

[12] The ICFES exam underwent a major reform in 2000 (Riehl, 2019), so from 2000–2002 some programs allowed students to apply with either old or new ICFES scores. Our regressions include dummies for program/cohort/ICFES-version triples because these variables define the relevant admission threshold.

[13] One exception is that the STEM earnings return decreases by about half for the most narrow bandwidth ($h = 15$). The change in this estimate is driven by high-ability applicants; for lower-ability applicants, there is a large positive STEM earnings return regardless of the bandwidth choice (see Appendix Figure A2).

[14] The first-stage 2SLS regression is equivalent to equation (2) with E_{ip} as the dependent variable.

by crossing the admission threshold.[15] As highlighted by our framework in Section 3, the coefficient β reflects the outcomes of both graduates and drop-outs. It also depends on the programs applicants chose when they were rejected, which we explore in Section 6.

We also estimate equation (3) with an indicator for graduating from the flagship program as the dependent variable, Y_{ip}. In this case, β measures the program graduation rate for compliers who are on the margin of admission.

4.2. **Identification assumptions and balance tests.** Our identification relies on the standard RD and instrumental variable assumptions. In our reduced-form model (2), the main identification assumption is that individuals near the threshold do not have perfect control over their admission status (Lee and Lemieux, 2010). Although students likely have an idea about the program's quota and standards, there is uncertainty in the final admission decision stemming from other applicants.

The appendix includes results that support this identification assumption. Appendix Table A5 shows little evidence that individual characteristics such as ICFES scores, gender, and socioeconomic background change discontinuously at the admission thresholds. Appendix Figure A1 shows that our sample also passes the McCrary (2008) density test, which indicates that the density of admission scores does not change discretely near the admission thresholds.

To interpret the 2SLS coefficient, β, as a mean return for compliers, we also rely on the standard Local Average Treatment Effect (LATE) assumptions (Angrist et al., 1996). Below we show that admission increased the probability of enrolling by 75 percentage points, which supports the assumption of instrument relevance. The exclusion restriction states that crossing the admission threshold affects individuals' outcomes only through the channel of flagship enrollment. This is a weaker assumption than that in papers with similar research designs that use peer characteristics (Abdulkadiroğlu et al., 2014) or degree completion (Kirkebøen et al., 2016) as the endogenous regressor. The monotonicity assumption requires that there are no applicants who would enroll in the flagship program if and only if they were just *below* the admission threshold, which is plausible in our setting.

4.3. **Mean returns.** Table 3 presents estimates of the effects of admission and enrollment in flagship programs. The sample for columns (A)–(B) includes applicants to all STEM programs, and columns (C)–(D) include applicants to other programs. Columns (A) and (C) display means of each dependent variable for applicants who were 1–5 positions below the admission thresholds. Columns (B) and (D) display estimates of the RD coefficients.

The first row of Table 3 shows that marginal STEM admits had slightly higher mean ability than admits to other programs. The mean applicant just below the STEM admission

[15] Non-compliers include above-threshold applicants who turned down their admission offer, and below-threshold applicants who enrolled in the same program by reapplying to a subsequent cohort.

thresholds scored at the 85^{th} percentile of the graduation-anchored ICFES exam, defined as in Section 1.3. The average marginal reject for other programs scored at the 82^{nd} percentile. In both program groups, we find no discontinuity in ability percentile at the admission thresholds, which supports our main RD identification assumption.

Panel A of Table 3 displays estimates of the reduced-form RD coefficient, θ, from equation (2), which represent the effects of flagship admission. Figure 2 also illustrates these reduced-form effects in RD graphs. The x-axis in each panel of Figure 2 is our running variable, x_{ip}, which is a student's position in their application pool normalized to zero at the threshold. Markers are means of the dependent variable listed in the panel title within eight unit bins of x_{ip}. We display separate means for STEM programs (red circles) and other programs (hollow triangles). Lines are predicted values for non-parametric regressions estimated separately above and below the threshold.

Admission to the flagship had similar effects on enrollment in STEM and other programs. Crossing the admission threshold increased the probability of flagship enrollment by 75 percentage points for STEM applicants, and 78 percentage points for other applicants. This large increase in flagship enrollment is also depicted in Panel A of Figure 2. Among marginal admits, the flagship's yield was about 90 percent in both program groups.

The flagship's STEM programs had significantly lower completion rates for marginal admits than in other programs. Crossing the admission threshold increased the probability of attaining a flagship degree by 26 percentage points in STEM programs, and 39 percentage points in other programs. This effect is illustrated in Panel B of Figure 2. Thus the degree attainment rate is nearly 14 percentage points lower for marginal admits to STEM programs.

Despite lower rates of degree completion, STEM admits had significantly higher earnings returns measured 13–17 years later. We find no significant effects on the likelihood of formal employment in either applicant group, defined as appearing in our earnings data (see also Panel C of Figure 2). Conditional on employment, admission to a STEM program raised applicants' monthly earnings in 2017 by about 10 percent. By contrast, admission to other flagship programs had no significant impact on average earnings. Panel D in Figure 2 shows a jump in mean log earnings at the admission thresholds for STEM applicants, and a small but insignificant decrease for other applicants.

Panel B of Table 3 displays 2SLS estimates of the graduation and earnings returns to flagship *enrollment*. These are estimates of the RD coefficients, β, from equation (3). The flagship graduation rate is equal to the ratio of the reduced-form coefficients for flagship completion and flagship enrollment. Similarly, the 2SLS earnings return equals the reduced-form earnings coefficient divided by the coefficient on flagship enrollment.

Marginal STEM admits had lower graduation rates but *higher* earnings returns than admits to other programs. The STEM graduation rate was only 34 percent, which is 15 percentage points lower than the mean graduation rate in other programs. Yet enrolling in a STEM program raised individuals' 2017 earnings by 13 percent on average, while there was no significant mean earnings returns to enrolling in other flagship programs.[16]

Our STEM graduation and earnings estimates are similar in magnitude to those in the literature. The 34 percent graduation rate for marginal STEM admits is comparable to persistence rates for less-prepared students at U.S. flagship schools. At UC Berkeley, for example, Arcidiacono et al. (2016) find that only 28 percent of minority students with an initial major in science earned a degree within five years. Using Norwegian data, Kirkebøen et al. (2016) find that the mean earnings premium of science and engineering degrees relative to individuals' next-choice fields is roughly 40 percent.[17] This is consistent with our mean earnings return of 13 percent given that our estimates reflect returns to STEM enrollment rather than degree attainment, and only one-third of enrollees complete a degree.

The results for other flagship programs also suggest that STEM degrees have a unique earnings benefit in Colombia. Below we show that only half of marginally-rejected STEM applicants enrolled in another STEM program. By contrast, marginal rejects from other flagship programs were *more* likely to enroll in other STEM programs. This may partly explain why the mean return to other flagship programs is close to zero.

4.4. **Returns by program selectivity.** Figure 3 shows how graduation and earnings returns vary with program selectivity. In both panels, the x-axis is a measure of the program's selectivity, defined as the mean ability percentile of applicants just below the admission thresholds. Selectivity varies substantially across programs in both STEM and other fields. On average, applicants who were barely rejected had ICFES scores well above the 90$^{\text{th}}$ percentile in highly selective programs like Medicine and Systems Engineering, and as low as the 65$^{\text{th}}$ percentile in Accounting and Topographical Engineering. The y-axis plots the program's graduation rate (Panel A) and earnings returns to enrollment (Panel B), which are β coefficients from separate estimations of equation (3) for each program.

Panel A shows that STEM programs have significantly lower graduation rates than similarly selective programs in other fields. There is a positive correlation between selectivity and graduation rates in both program groups, but completion rates are systematically lower in STEM programs. For example, the graduation rate for marginal admits to the flagship's selective Medicine program was over 80 percent, while only 33 percent of marginal enrollees in Systems Engineering completed the program. Graduation rates for marginal admits were as low as 20 percent in less selective STEM programs like Physics and Statistics.

[16] We do not find evidence of significant earnings returns to any of the non-STEM faculty areas in Table 2, although sample sizes are small.

[17] In Kirkebøen et al. (2016), the average payoffs to science and engineering degrees are roughly 23,000 U.S. dollars (Figure IX), and mean earnings in their sample is 56,000 U.S. dollars (Table III).

This pattern is reversed for earnings, with STEM programs delivering higher earnings returns than other similarly selective programs. These returns are noisily estimated for individual programs, but Panel B of Figure 3 shows that STEM programs had systematically higher returns conditional on selectivity. For example, the point estimates suggest large earnings returns to Systems, Chemical, and Mechanical Engineering despite their low graduation rates, while returns are actually negative (but insignificant) in Medicine and Odontology.[18]

The results in this section show that mean graduation rates and earnings returns were inversely related across flagship programs. Next we ask if a similar pattern holds *within* programs for applicants with different levels of academic preparation.

5. Heterogeneity in STEM returns

This section uses two different empirical strategies to examine how returns to STEM enrollment varied with students' academic preparation. We first estimate heterogeneity in the RD coefficients using additional measures of applicant ability. We then exploit variation in the flagship's admission quotas in an RD differences-in-differences design. In both approaches, we find that less-prepared applicants were less likely to complete the flagship's STEM programs, but they had *higher* earnings returns to enrolling.

5.1. **Heterogeneity in RD coefficients.** Our first approach to estimating heterogeneity in STEM returns exploits variation that arises because the flagship's admission scores are a noisy measure of ability. These admission scores are a weighted average of applicants' ICFES subject scores, which included up to nine different tests during the period of our data. The weights varied across programs, and some ICFES scores received zero weight in subjects that were less related to the major. As a result, applicants with similar admission scores varied in their ability as defined by other averages of the ICFES subject scores.

Figure 4 shows how our ability measure varied for applicants with similar admission scores. The x-axis in each panel is an applicant's position in their application pool normalized to zero at the threshold, as in Figure 2. The y-axis is our ability measure, which is an applicant's percentile in the national distribution of graduation-anchored ICFES scores. We demean ability within each application pool to focus on variation among students who faced the same admission threshold. The solid line depicts the median of our demeaned ability measure in eight unit bins of admission rank. Dashed lines depict the 10^{th} and 90^{th} percentiles. Panel A includes STEM applicants, while Panel B includes applicants to other programs.

Applicants near the same admission threshold varied significantly in graduation-anchored ability. Among marginal STEM admits (Panel A), the median student scored six percentile

[18] We note that Economics is similar to STEM programs in that it has a low graduation rate for marginal admits (Panel A of Figure 3). We find no evidence of a positive earnings return to Economics (Panel B), but this estimate is imprecise.

points higher on the ICFES exam than the average student in their application pool. The 90–10 range for marginal admits includes students who scored 18 percentile points above the mean and 10 percentile points below the mean. The variation in ability is even larger among marginal admits to other flagship programs (Panel B).

We exploit this variation by dividing our sample into groups of above- and below-median ability conditional on admission rank. Specifically, we regress applicants' graduation-anchored ability on their admission rank with application pool dummies, and take the residuals from this regression. We define "lower-ability" applicants as individuals whose residual ability is below the median in their application pool. "High-ability" applicants are those with above-median residuals from this regression.

Table 4 shows how the returns to flagship programs differed in the lower- and high-ability samples.[19] Panel A displays the reduced-form RD coefficients, θ, from equation (2). Panel B shows the 2SLS estimates, β, from equation (3). We first focus on the results from STEM applicants, which appear in columns (A)–(C). Columns (A) and (B) display these RD coefficients from separate regressions for the lower- and high-ability samples. Column (C) shows the p value from an F test for equality of the two coefficients.

The first row of Table 4 shows that our two samples differ by more than 15 percentile points in mean ability near the admission thresholds. This row displays mean graduation-anchored ability for applicants 1–5 positions below the admission thresholds. On average, marginal STEM rejects scored at the 92[nd] percentile of the ICFES exam in our high-ability sample, and at the 77[th] percentile in our lower-ability sample. Thus the two samples differ significantly in propensity to graduate as predicted by ICFES scores.

The flagship enrollment rate was similar for lower- and high-ability marginal admits, but lower-ability admits were significantly less likely to complete the STEM programs. Admission to a flagship STEM program increased the probability of enrollment by about 75 percentage points in both samples. However, admission increased the likelihood of earning a flagship STEM degree by 30 percentage points in the high-ability sample, and by only 20 percentage points in the lower-ability sample. Panel B shows that the graduation rate for marginal STEM admits was 13 percentage points lower in the less able sample, and we reject equality of the two coefficients.

Despite lower completion rates, lower-ability STEM admits had *higher* earnings returns to enrolling. We find no significant effects on formal employment in either sample, and the two point estimates are not statistically different. By contrast, the earnings return to the STEM programs is substantially larger for less-prepared applicants. Admission to a flagship STEM program raised 2017 earnings by about 20 percent for lower-ability applicants, while it had no significant impact on earnings for high-ability applicants. Panel B shows that

[19] Appendix Figure A2 presents RD graphs that correspond to the results in Table 4.

lower-ability admits had large returns to STEM enrollment, and we can reject that this coefficient is equal to that of high-ability enrollees at the 10 percent level.

Columns (D)–(F) show that graduation rates and earnings returns varied less with ability in other programs. These columns are identical to columns (A)–(C), but the sample includes applicants to all non-STEM programs. The graduation rate for these marginal admits was eight percentage points lower for less able applicants than for high-ability applicants (45 percent vs. 53 percent). This gap is roughly half the magnitude of that for STEM applicants, despite a larger difference in mean ability between the two samples. The point estimates for formal employment and earnings are not statistically different from zero in either sample.

The results in Table 4 are robust to different measures of ability. Appendix Table A6 shows that our findings are similar when we define lower- and high-ability samples using the simple average of ICFES subject scores rather than with graduation anchoring. The 2000 overhaul of the ICFES exam (see Section 4.1) may also have impacted our ability measure by changing the scores' predictive power or the flagship's admission criteria. We find similar results when we restrict our sample to applicants with post-2000 ICFES scores, so that we define high- and lower-ability using a single version of the exam (see Appendix Table A7).

5.2. **Variation in admission quotas.** Our second approach to estimating heterogeneity in STEM returns exploits variation across cohorts in the flagship's admission quotas. This section describes the sources of this quota variation.

Table 5 shows the number of students who were admitted to the flagship's STEM programs by semester of application. The flagship is similar to most Colombian colleges in that it frequently offers cohorts that begin in either August or January. The number of admits per semester and program typically ranged from 45–65 students during the period of our data. Most of the 18 STEM programs in our sample had a relatively stable number of admits per cohort during this time, as shown in the last row of Table 5.

In six STEM programs, however, admission quotas increased significantly in certain cohorts. Larger quotas occurred for two reasons. First, the Biology and Systems Engineering programs changed their desired cohort size in some years. Table 5 shows that the Biology program admitted cohorts of 80–100 students each fall from 1999–2001. In January 2002, it began admitting cohorts of roughly 40–60 students each semester (except for another large cohort in August 2002). Similarly, Systems Engineering typically had cohorts of about 60 students, but in August 2001 it admitted more than 120 students.

The second source of quota variation came from changes in how the flagship conducted admissions. Programs typically would use separate admission pools when they offered both fall and spring cohorts. In certain years, however, four engineering programs "tracked" applicants into fall and spring cohorts all at once. Under tracking admissions, roughly the top 60 students in the application pool were admitted to an August cohort, while the next 60 students were admitted to a January cohort. Table 5 shows that the Chemical, Electrical, and Electronic Engineering programs used tracking admissions in August 2000, while Mechanical Engineering did so in August 2001.[20] In each case, more than 120 students were admitted in the tracking year, with the lower half starting at the flagship in the following January. Thus tracking caused the number of admits to roughly double.

If ability distributions are similar across cohorts, then the mean ability of marginal admits is decreasing in the size of the admission quota. This is likely to be true in Colombia because there is considerable variation in the timing at which students apply to college. During this time period, less than 20 percent of Colombians who attended college enrolled immediately after high school graduation, and the average gap between high school and college was nearly three years (de Roux and Riehl, 2019a). In our sample, the standard deviation of ages within each application pool is 2.6 years. Thus applicant characteristics do not differ systematically between the fall and spring, and larger quotas would tend to decrease the mean ability of marginal admits. We confirm this empirical pattern in our results below.

5.3. **RD differences-in-differences specification.** Our second empirical specification combines our RD model with a differences-in-differences regression that exploits variation in quotas across programs and cohorts. We begin by describing our reduced-form specification, and then discuss our 2SLS model.

The intuition for this specification comes from a two-step estimation procedure. In the first step, we estimate an RD coefficient for each application pool using the following equation:

$$Y_{ip} = \theta_p D_{ip} + \alpha_p x_{ip} + \psi_p D_{ip} x_{ip} + \gamma_p + \epsilon_{ip} \quad \text{if } |x_{ip}| \leq h. \tag{4}$$

This equation is similar to our reduced-form RD regression (2), but we interact all covariates with dummies for application pools, p (i.e., program/cohort pairs). As in equation (2), Y_{ip} is an outcome for individual i, x_{ip} is the applicant's admission rank, and D_{ip} is an indicator for having an admission score above the threshold. The coefficients of interest, θ_p, are the effects of flagship admission for each application pool, p.

The second step relates these RD coefficients to variation in admission quotas across programs and cohorts. In this step, we notate application pools, p, by the combination of program, m, and application cohort, t. The second-step regression uses the dependent variable $\theta_p = \theta_{mt}$ in a standard differences-in-differences model:

$$\theta_{mt} = \gamma_m + \gamma_t + \pi L_{mt} + \upsilon_{mt}. \tag{5}$$

[20] de Roux and Riehl (2019b) use this tracking admissions system to isolate the effects of mean classmate ability on an individual's returns to college. Materials Engineering also used this form of admissions in 2001, but this cohort is not in our sample because tracking caused all applicants to be admitted.

Equation (5) includes fixed effects for programs, γ_m, and semester of application, γ_t. These dummies control for fixed differences in graduation rates or earnings across programs, as well as aggregate changes in these outcomes over time.

Our variable of interest, L_{mt}, is an indicator for admission pools with large quotas. We define this variable in two ways. First, we define L_{mt} to be a binary variable for program-cohorts with unusually large quotas, as indicated by the bold numbers in Table 5. Specifically, $L_{mt} = 1$ for the large cohorts in the six programs listed in Table 5, and $L_{mt} = 0$ for all other cohorts of these programs and for all other STEM programs. Second, we define L_{mt} as the number of admitted students in each program-cohort, as indicated by the integers in all cells of Table 5. In this case, L_{mt} is a positive number for all program/cohort pairs, but most of the variation is driven by the unusually large quotas.

The coefficient of interest in equation (5) is π, which shows how the returns to flagship admission, θ_{mt}, changed when the quota increased. To show how this quota variation affected the ability of marginal admits, we also estimate versions of equation (5) in which the dependent variable is the mean ability of applicants just below the admission threshold. In this case, π measures how quota increases affected the mean ability of marginal rejects. For the results below, we estimate π in a single-step regression, which we derive by plugging equation (5) into equation (4). In this analysis we cluster standard errors at the program/cohort level, which is the level of variation in our treatment variable, L_{mt}.

We also estimate a 2SLS version of our RD differences-in-differences specification. Our 2SLS model is similar to our reduced-form specification, but we replace the above-threshold dummy, D_{ip}, with an indicator for flagship enrollment, E_{ip}, in equation (4). We then instrument for E_{ip} with a full set of interactions between D_{ip} and application pool dummies. In these 2SLS regressions, the coefficient π shows how the returns to STEM *enrollment* changed in cohorts with larger quotas.[21]

Identification of the coefficient π relies on the RD assumptions discussed in Section 4.2, as well as the standard differences-in-differences assumption of parallel trends. Our focus is on how STEM returns vary with ability, and thus our variable of interest, L_{mt}, is explicitly correlated with the characteristics of marginal admits. A violation of parallel trends would arise, for example, if macroeconomic shocks correlated with L_{mt} affected aggregate STEM

[21] Formally, our single-step 2SLS model is:

$$E_{ip} = \theta_p D_{ip} + \alpha_p x_{ip} + \psi_p D_{ip} x_{ip} + \gamma_p + \epsilon_{ip} \quad \text{if } |x_{ip}| \leq h$$
$$Y_{ip} = \big(\tilde{\gamma}_m + \tilde{\gamma}_t + \tilde{\pi} L_{mt}\big) E_{ip} + \tilde{\alpha}_p x_{ip} + \tilde{\psi}_p D_{ip} x_{ip} + \tilde{\gamma}_p + \tilde{\epsilon}_{ip} \quad \text{if } |x_{ip}| \leq h.$$

The first stage regression is equivalent to equation (4) with an indicator for enrolling in the flagship program, E_{ip}, as the dependent variable. The second stage regression is equivalent to our single-step reduced-form regression with D_{ip} replaced by E_{ip}.

returns in the economy. This is less likely given the haphazard timing of large quotas and the fact that our sample includes similar programs with limited quota variation (Table 5).

In this specification, our estimates may also be affected by other ways in which quota variation altered individuals' education at the flagship. For example, larger quotas may have impacted class sizes and peer characteristics, and "tracking" admissions affected the timing at which individuals enrolled. We discuss how these potential mechanisms affect the interpretation of our results in Section 6.1.

5.4. **Effects of quota variation on STEM returns.** Table 6 shows how variation in the flagship's quotas affected the ability of marginal STEM admits and their returns. Columns (A) and (B) display RD coefficients from equations (2)–(3) for applicants to the six programs in Table 5 with significant quota variation over time. Column (A) presents estimates for large-quota cohorts of these programs—indicated by the bold numbers in Table 5. Column (B) shows estimates for the small-quota cohorts. We weight observations in these regressions so that the distribution of programs is similar in the two columns.[22] Column (C) shows the p value from an F test that the coefficients are equal in the two cohort groups. Panel A presents reduced-form RD coefficients from equation (2), and Panel B displays 2SLS RD coefficients from equation (3).

Columns (D)–(E) in Table 6 present estimates of π from our RD differences-in-differences specification (4)–(5). Column (D) shows results using our binary measure of large quotas, L_{mt}, which is an indicator for the bolded cohorts in Table 5. Column (E) displays estimates in which L_{mt} is equal to the total number of admits in each program-cohort. We divide this measure of L_{mt} by 60, which is roughly the mean difference between large- and small-quota cohort sizes (see Table 5). These regressions include applicants to all STEM programs, and thus programs with limited quota variation are the "control group." Panel A presents reduced-form effects that show how large quotas affected the returns to STEM admission. Panel B presents 2SLS estimates of the returns to STEM enrollment.

The first row of Table 6 shows that the mean ability of marginal admits was significantly lower in application pools with large quotas. These estimates show how quota variation affected the mean graduation-anchored ability of applicants 1–5 positions below the admission thresholds. Column (A)–(B) show that these marginal rejects scored at the 83$^{\text{rd}}$ percentile of the IFES exam in large-quota cohorts, and at the 92$^{\text{nd}}$ percentile in small-quota cohorts. In our differences-in-differences specification in columns (D)–(E), the larger quotas are associated with an eight percentile point reduction in the ability of marginal admits using both measures of L_{mt}.

[22] These weights make the difference between the RD estimates in columns (A) and (B) roughly comparable to the RD differences-in-differences coefficient in column (D). See the notes to Table 6 for details.

Marginal admits in cohorts with larger quotas were significantly less likely to complete the flagship's STEM programs if they enrolled. Panel A shows that the flagship enrollment rate for marginal admits was significantly higher in large-quota cohorts (78 percent) than in cohorts with smaller quotas (61 percent). Yet admission raised the probability of completing a flagship STEM degree by about 27 percentage points in both groups. Thus the flagship graduation rate for marginal admits was significantly lower in cohorts with large quotas (34 percent) than in those with smaller quotas (45 percent), as shown in Panel B. Larger quotas are associated with a 12–13 percentage point reduction in the graduation rate in our RD differences-in-differences specifications (columns (D)–(E)).

Despite lower graduation rates, the earnings return to STEM enrollment was significantly *higher* for marginal admits in cohorts with large quotas. We find no significant effects on formal sector employment rates for either large- or small-quota cohorts. The earnings return to admission is 0.41 log points in cohorts with large quotas (column (A)), and it is only seven percent in cohorts with smaller quotas (column (B)). The last row of Panel B shows that the quota increase is associated with more than a 40 percent increase in the earnings return to STEM enrollment in our RD differences-in-differences specification (column (D)), although this coefficient has a large standard error. Column (E) shows similar results using the number of admitted students as our measure of L_{mt}, but the effects are smaller in magnitude and are not statistically significant.

In sum, Table 6 shows that lower-ability admits were less likely to complete the flagship's STEM programs, but they had higher earnings returns to enrollment.[23] These findings are consistent with the RD heterogeneity results in Table 4. Appendix Table A9 shows a similar pattern of results using a differences-in-differences specification that exploits variation in admission rates across all ability levels, rather than just at the admission thresholds.[24]

The inverse relationship between graduation rates and earnings returns across ability levels also appears to be unique to the flagship's STEM programs. Appendix Tables A10–A11 replicate the analysis in Tables 5–6 for other programs at the flagship. This analysis is underpowered because there is less variation in quotas for other programs. However, graduation rates are less related to the mean ability of marginal admits in other programs, and there is little evidence that earnings returns were higher in cohorts with large quotas.

6. Mechanisms

This section examines mechanisms for why less-prepared applicants had higher earnings returns to enrolling in the flagship's STEM programs despite lower graduation rates. We

[23] Our findings are similar when we include only applicants with post-2000 ICFES scores, although the RD differences-in-differences earnings effects are smaller and statistically insignificant (see Appendix Table A8).
[24] We thank an anonymous referee for suggesting this analysis.

first discuss the potential role of peer and class size effects in our analysis. We then consider the three mechanisms highlighted by our framework in Section 3. We find little support for the hypothesis that flagship drop-outs had large returns to STEM enrollment. Instead, we present evidence that our results are driven by heterogeneity in both the program choices of rejected applicants and in the value added of a STEM degree.

6.1. **Peer and class size effects.** The results in Section 5 suggest that the returns to selective STEM programs vary with students' pre-college ability. However, our identification strategies are also associated with variation in other factors that could influence an individual's returns, such as peer composition (Sacerdote, 2001) or class size (Angrist and Lavy, 1999). This section discusses how these other potential mechanisms affect the interpretation of our results.

Our two identification strategies in Section 5 have different sets of potential mechanisms. Section 5.1 estimated heterogeneity in RD coefficients using variation in ability *within* each application pool. Identification comes from high- and lower-ability students who were admitted to the same programs and cohorts, and who therefore mostly sat in the same classrooms. Thus this strategy estimates how STEM returns vary with ability holding peer composition fixed. These estimates are most relevant to debates on admission policies like affirmative action that affect a relatively small number of students. At many U.S. colleges, for example, eliminating affirmative action may cause a few more-prepared students to enroll in selective STEM programs, but it would not have a dramatic effect on peer composition.

Conversely, our RD differences-in-differences strategy in Sections 5.3–5.4 includes changes in class size and peer composition. In this approach, identification comes from programs that roughly doubled their admission quotas. This reduced the ability of marginal admits, but it also decreased *average* ability in the affected cohorts and, in some cases, led to larger class sizes.[25] These estimates are most relevant to policies that expand enrollment in selective STEM programs through increases in university funding. The effects of such policies depend on how less-prepared students perform in programs with lower mean ability and potentially larger classes. However, our estimates do not capture general equilibrium effects that might arise from very large increases in STEM enrollment, such as those in Bianchi (2020). For example, earnings returns might be lower if policy changes were large enough to affect the relative supply of STEM workers in the labor market.

[25] We do not have data on class size, but anecdotally the flagship accommodated larger admission quotas by expanding enrollment in some courses and adding new sections in others. Changes in class size were likely most relevant in the Biology and Systems Engineering programs, in which all students in the larger cohorts enrolled at the same time (see Table 5). Conversely, changes in peer quality were likely larger in the four programs that tracked students into separate fall and spring cohorts.

Our RD differences-in-differences strategy also relies partially on variation in the time at which students enrolled in the flagship. In the four programs with tracking admissions, marginal admits in the large cohorts had to wait approximately five months before enrolling (see Table 5). Thus variation in returns between large- and small-quota cohorts could be partly affected by timing mechanisms such as learning decay (Cooper et al., 1996) or age-at-enrollment effects (Bedard and Dhuey, 2006). de Roux and Riehl (2019b) test for and find little evidence that enrollment timing affected graduation rates and earnings using a similar sample to that in this paper, but we cannot completely rule out this mechanism.

Our two identification strategies yield similar results despite these differences in potential mechanisms, which suggests that heterogeneity in STEM returns is primarily attributable to variation in individuals' own ability. We find similar gaps in graduation rates between high- and lower-ability students in our RD heterogeneity and RD differences-in-differences analyses (Tables 4 and 6). The variation in earnings returns is slightly larger in Table 6, but we cannot statistically reject that the gaps are equal given our large standard errors. These similarities suggest that variation in individuals' academic preparation is a more important factor in our results than variation in class size or peer composition.

The rest of this section considers potential reasons why less-prepared applicants had larger returns to STEM enrollment than more able applicants, despite their lower graduation rates.

6.2. **Returns for flagship drop-outs.** One potential explanation for our results is that there may be benefits to enrolling in a selective STEM program even if students do not graduate. Our framework in Section 3 illustrates this possibility by allowing for a value added of program *enrollment*. For example, many Colombian students work part-time while they are in school, and STEM enrollment may give them access to certain jobs or internships. Enrollment may also provide access to peer networks that help individuals' careers even if they do not graduate.

To examine this hypothesis, Figure 5 displays earnings outcomes for students who did not complete the flagship STEM program. For this figure, we divide our sample of STEM applicants into two groups: those who earned a college degree, and those who did not. In both panels, the x-axis shows individuals' position in their application pool normalized to zero at the threshold, as in Figure 2. Since our administrative data include nearly all Colombian colleges, our sample includes graduates and non-graduates both above and below the flagship's admission thresholds.

Panel A shows that college graduates and non-graduates had similar mean ability on both sides of the threshold, which suggests that flagship admission did not significantly affect selection into these two samples. The y-axis in this panel is an applicant's graduation-anchored ability percentile. Mean ability is at least five percentile points higher for graduates than for non-graduates at all admission ranks. However, there is no evidence that ability changes discontinuously in either group at the flagship's admission thresholds. Although these two samples are defined by an endogenous outcome, this result suggests that selection into the graduate sample is similar for both flagship admits and rejects.

Panel B shows that the STEM earnings return appears to be driven by graduates rather than drop-outs. The y-axis in this panel is log monthly earnings in 2017. Mean earnings are similar for marginal admits and marginal rejects who did not complete college. Both groups have mean monthly earnings of about 14.1 log points, which equates to an annual salary of 4,700 U.S. dollars. Among college graduates, mean earnings increase by over 20 percent at the thresholds. Marginal admits who completed a degree made almost twice as much on average as those who did not, with a mean annual salary of about 8,500 U.S. dollars.

In sum, Figure 5 shows that STEM applicants who did not earn college degrees were similar in both mean ability and mean earnings on both sides of the admission thresholds. Appendix Figure A3 shows similar results in a sample of lower-ability applicants. These findings are hard to rationalize with large returns for the less-prepared applicants who did not earn a STEM degree.[26] Thus the inverse relationship between graduation rates and earnings returns across ability levels is unlikely to be driven by outcomes for drop-outs.

6.3. **Counterfactual program choices.** Another potential explanation for our results is that more- and less-prepared applicants may have chosen different programs when they were not admitted. Our framework in Section 3 shows that the returns to a selective STEM program depend in part on the skill an individual would gain in their next-choice program. Less-prepared students may have had larger earnings returns to the flagship's STEM programs if they enrolled in programs with lower value added when they were rejected.

To test this hypothesis, Table 7 shows how the counterfactual programs of flagship applicants varied by ability. This table uses our administrative higher education records to measure enrollment and graduation outcomes at nearly all Colombian colleges. We divide flagship STEM applicants into lower- and high-ability samples using the same procedure as for Table 4. Column (A) shows the mean of each outcome for lower-ability applicants who were 1–5 positions below the flagship's admission thresholds. Column (B) displays the reduced-form RD coefficient, θ, from estimating equation (2) in the lower-ability sample. Columns (C)–(D) present the same results for the high-ability sample. Column (E) displays the p value from an F test of equality of the RD coefficients in columns (B) and (D).

Lower-ability applicants were more likely to forgo college altogether if they were not admitted to the flagship. The first row of Table 7 shows that 77 percent of lower-ability applicants

[26] We also find that flagship admission increased the total time in college by only 0.8 years on average across all applicants. Thus flagship drop-outs did not have significantly more schooling than their below-threshold counterparts, which may explain why they did not experience a significant earnings return.

who were marginally rejected enrolled in any college program. Flagship admission raised the probability that less able applicants enrolled in college by 17 percentage points. Among high-ability applicants, 90 percent of marginal rejects attended college, and admission raised the college enrollment rate by only eight percentage points. These results are consistent with constraints in supply of desirable programs, as even the "lower-ability" applicants in our sample scored at the 77$^{\text{th}}$ percentile of the ICFES exam on average.

Less-prepared applicants were also less likely to enroll in any STEM program if they were rejected by the flagship. The second row of Table 7 shows that 48 percent of lower-ability applicants just below the threshold enrolled in another STEM program, as compared with 62 percent of high-ability applicants. Flagship admission had a big effect on the likelihood of choosing a STEM program in both groups, but the effect was about 10 percentage points larger for less-prepared applicants. Much of this difference is driven by the fact that lower-ability applicants were less likely to attend any college if they were rejected, but they were also less likely to choose a STEM program conditional on college attendance. Among marginal rejects who enrolled in any college, the fraction who picked a STEM program was 62 percent for lower-ability applicants and 69 percent for high-ability applicants.

The last two rows of Table 7 show that flagship admission had similar effects on overall degree attainment for high- and lower-ability applicants. Although less-prepared applicants had low flagship completion rates, admission to the flagship still raised their probability of earning a college degree by eight percentage points, and it increased their likelihood of earning any STEM degree by 16 percentage points. The effects of flagship admission on degree attainment are similar in magnitude for high-ability applicants, despite the fact that admission had less impact on their enrollment choices.

The results in Table 7 can partially explain the heterogeneity in our earnings results. Degree attainment effects are similar across ability levels because the gap in flagship completion rates was offset by the gap in counterfactual enrollment choices. Thus the difference in earnings returns between lower- and high-ability applicants would likely have been smaller if they had similar counterfactual enrollment choices. We illustrate this in Appendix Table A12 by showing how flagship admission affected the mean earnings in applicants' college programs, which we measure using our national administrative data. For lower-ability students, admission to a flagship STEM program increased mean earnings in an individuals' programs by about 12 percent. The effect of admission on mean earnings was only six percent for high-ability applicants.

Yet next-choice programs cannot explain why less-prepared applicants had *larger* earnings returns because flagship admission had similar effects on degree attainment. Thus we turn to our final potential mechanism.

6.4. **Heterogeneity in value added.** From Section 3, the other primary explanation for our results is that the value added of a flagship STEM degree is decreasing in ability. Less-prepared students may have had to work harder to complete the program, and thus learned more along the way. Value added could also decrease with ability for reasons unrelated to skill accumulation. A STEM degree may provide access to high-paying industries or firms in a labor market with hiring frictions. In this case, lower-ability students may gain more from a degree if they have worse job opportunities without one.

Figure 6 presents evidence on this hypothesis by showing how a flagship STEM degree affects the relationship between earnings and ability. This figure plots log monthly earnings in 2017 (y-axis) against applicants' graduation-anchored ability percentile (x-axis). The markers depict mean earnings in two-percentile bins for different groups of students. The solid line includes all Colombian high school graduates in our administrative data who did not earn a college degree. The hollow triangles include flagship STEM applicants who did not earn a degree from that program. Lastly, the red circles include applicants who earned a flagship STEM degree.

The results in Figure 6 are consistent with the hypothesis that the value added of a flagship STEM degree decreases with ability. The solid line shows that there is a strong relationship between earnings and ability in the population of high school graduates without college degrees. Within this sample, individuals at the top of the ability distribution earn about 66 percent more than those at the 60$^{\text{th}}$ percentile. This pattern is similar for flagship STEM applicants who did not complete the program. But the earnings/ability relationship is much weaker for applicants who completed a flagship STEM degree. Among STEM degree recipients, a ten percentile increase in ability is associated with only a four percent increase in mean earnings. Further, graduates below the 80$^{\text{th}}$ percentile of ability had mean earnings that were similar to those of higher-ability graduates.

Figure 6 suggests that a flagship STEM degree may help less-prepared students "catch up" to their more able peers. This result is consistent with the regressions in Table 1, which show a flat earnings/ability gradient using all STEM programs in Colombia. These analyses do not conclusively show that STEM degrees caused the earnings/ability gradient to decline because lower-ability students who completed a degree are likely to be positively selected.[27] As a partial test of this possibility, we replicate Figure 6 using residuals from an earnings regression that includes controls for individuals' high schools, gender, age, and socioeconomic backgrounds. Appendix Figure A4 shows a similar pattern of results using earnings residuals, which suggests that the flat gradient for STEM graduates is not driven by differences in the observable characteristics of high- and lower-ability individuals.

[27] On the other hand, high-ability students who did not earn degrees are likely negatively selected.

We cannot rule out the possibility that the findings in Figure 1 reflect differences in unobservable applicant traits. However, the change in the earnings/ability relationship is large enough to explain the gap in returns between lower- and high-ability STEM applicants, and we cannot fully explain this gap through other mechanisms. Thus we conclude that less-prepared applicants likely had an especially large value added to a flagship STEM degree.

7. Conclusion

On many college campuses, STEM programs have a reputation for "weeding out" underperforming students through low grades. At University of California schools, Arcidiacono et al. (2016) show that students with relatively less academic preparation were more likely to switch out of STEM majors or drop out of college altogether. The authors' results suggest that less-prepared students would have been more likely to earn STEM degrees if they had enrolled at less selective colleges.

This paper has shown evidence consistent with this finding using data from a flagship university in Colombia. In two different empirical designs, we found that less-prepared students were significantly less likely to complete the selective STEM programs. At the university in our analysis, the graduation rate for marginal STEM admits was only 34 percent, and degree attainment was even less common among enrollees with below-median test scores.

On the other hand, our results show that raising admission standards may not necessarily increase the likelihood that less-prepared students attain STEM degrees. Using national enrollment data, we found that less-prepared applicants were less likely to enroll in another STEM program when they were rejected, and also less likely to enroll in college at all. Despite their lower graduation rates, less-prepared students were still more likely to earn a STEM degree if they were admitted to the flagship. This may be partly attributable to constraints in the supply of high-quality STEM programs in Colombia. But this result shows that it is important to consider counterfactual enrollment choices in the debate on how admission preferences affect STEM degree attainment.

Our paper also shows why less-prepared students might choose to enroll in selective STEM programs despite lower completion rates. Using data on earnings roughly 15 years after flagship application, we found that returns to STEM enrollment were *higher* for less-prepared students than for their more able peers. Our analysis of mechanisms suggests that this was driven by especially large gains for the individuals who successfully completed a degree. We found similar patterns in our descriptive analysis of earnings premiums by ability and degree field using national administrative data. Our results are thus consistent with a standard human capital model in which less-prepared students enroll in programs that maximize their expected returns, even if they are aware of the low probability of success.

Our findings suggest that policies that attract students to enroll in selective STEM programs can yield large earnings gains, even if these students are less-prepared on average. An important caveat is that many individuals may not experience these gains if they do not complete a degree. This highlights the importance of other initiatives that help students to develop STEM skills at younger ages.

References

Abdulkadiroğlu, A., J. Angrist, and P. Pathak (2014). The elite illusion: Achievement effects at Boston and New York exam schools. *Econometrica 82*(1), 137–196.

Altonji, J. G. (1993). The demand for and return to education when education outcomes are uncertain. *Journal of Labor Economics 11*(1, Part 1), 48–83.

Altonji, J. G., P. Arcidiacono, and A. Maurel (2016). The analysis of field choice in college and graduate school: Determinants and wage effects. In *Handbook of the Economics of Education*, Volume 5, pp. 305–396. Elsevier.

Altonji, J. G., E. Blom, and C. Meghir (2012). Heterogeneity in human capital investments: High school curriculum, college major, and careers. *Annual Review of Econonomics 4*(1), 185–223.

Anelli, M. (2018). The returns to elite university education: A quasi-experimental analysis. Working Paper.

Angrist, J. D., G. W. Imbens, and D. B. Rubin (1996). Identification of causal effects using instrumental variables. *Journal of the American Statistical Association 91*(434), 444–455.

Angrist, J. D. and V. Lavy (1999). Using Maimonides' rule to estimate the effect of class size on scholastic achievement. *The Quarterly Journal of Economics 114*(2), 533–575.

Angrist, J. D. and M. Rokkanen (2015). Wanna get away? regression discontinuity estimation of exam school effects away from the cutoff. *Journal of the American Statistical Association 110*(512), 1331–1344.

Arcidiacono, P. (2004). Ability sorting and the returns to college major. *Journal of Econometrics 121*(1), 343–375.

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

Arcidiacono, P., E. M. Aucejo, and K. Spenner (2012). What happens after enrollment? an analysis of the time path of racial differences in GPA and major choice. *IZA Journal of Labor Economics 1*(1), 1.

Bedard, K. and E. Dhuey (2006). The persistence of early childhood maturity: International evidence of long-run age effects. *The Quarterly Journal of Economics 121*(4), 1437–1472.

Bianchi, N. (2020). The indirect effects of educational expansions: Evidence from a large enrollment increase in university majors. *Journal of Labor Economics 38*(3).

Bianchi, N. and M. Giorcelli (2019). Scientific education and innovation: from technical diplomas to university STEM degrees. *Journal of the European Economic Association.*

Calonico, S., M. D. Cattaneo, and R. Titiunik (2014). Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica 82*(6), 2295–2326.

Canaan, S. and P. Mouganie (2018). Returns to education quality for low-skilled students: Evidence from a discontinuity. *Journal of Labor Economics 36*(2), 395–436.

Carnevale, A. P., N. Smith, and M. Melton (2011). STEM: Science technology engineering mathematics. *Georgetown University Center on Education and the Workforce.*

Chetty, R., J. N. Friedman, E. Saez, N. Turner, and D. Yagan (2017). Mobility report cards: The role of colleges in intergenerational mobility. Technical report, National Bureau of Economic Research.

Cooper, H., B. Nye, K. Charlton, J. Lindsay, and S. Greathouse (1996). The effects of summer vacation on achievement test scores: A narrative and meta-analytic review. *Review of*

Educational Research 66(3), 227–268.

Cunha, F., J. J. Heckman, and S. M. Schennach (2010). Estimating the technology of cognitive and noncognitive skill formation. *Econometrica 78*(3), 883–931.

de Roux, N. and E. Riehl (2019a). Disrupted academic careers: The returns to time off after high school. Working Paper.

de Roux, N. and E. Riehl (2019b). Isolating peer effects in the returns to college selectivity. Working Paper.

Deming, D. and K. Noray (2019). STEM careers and the changing skill requirements of work. HKS Working Paper No. RWP19-025.

Duflo, E., P. Dupas, and M. Kremer (2011). Peer effects, teacher incentives, and the impact of tracking: Evidence from a randomized evaluation in Kenya. *American Economic Review 101*(5), 1739–1774.

Goodman, J., M. Hurwitz, and J. Smith (2017). Access to 4-year public colleges and degree completion. *Journal of Labor Economics 35*(3), 829–867.

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. National Bureau of Economic Research Working Paper 19241.

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

Holdren, J. P. (2013). Federal science, technology, engineering, and mathematics (stem) education: 5-year strategic plan. Technical report, Committee on STEM Education of the National Science and Technology Council.

Hoxby, C. M. and C. Avery (2013). Missing one-offs: The hidden supply of high-achieving, low-income students. Brookings Papers on Economic Activity.

Kinsler, J. and R. Pavan (2015). The specificity of general human capital: Evidence from college major choice. *Journal of Labor Economics 33*(4), 933–972.

Kirkebøen, L., E. Leuven, and M. Mogstad (2016). Field of study, earnings, and self-selection. *The Quarterly Journal of Economics 131*(3), 1057–1111.

Lee, D. S. and T. Lemieux (2010). Regression discontinuity designs in economics. *Journal of Economic Literature 48*, 281–355.

Machado, C. and C. Szerman (2017). Centralized admissions and the student-college match. SSRN working paper.

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics 142*(2), 698–714.

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

Pallais, A. (2015). Small differences that matter: Mistakes in applying to college. *Journal of Labor Economics 33*(2), 493–520.

Peri, G., K. Shih, and C. Sparber (2015). STEM workers, H-1B visas, and productivity in US cities. *Journal of Labor Economics 33*(S1), S225–S255.

Pop-Eleches, C. and M. Urquiola (2013). Going to a better school: Effects and behavioral responses. *American Economic Review 103*(4), 1289–1324.

Riehl, E. (2019). Fairness in college admission exams: From test score gaps to earnings inequality. Working Paper.

Saavedra, J. E. (2009). The returns to college quality: A regression discontinuity analysis. Harvard University.

Sabot, R. and J. Wakeman-Linn (1991). Grade inflation and course choice. *Journal of Economic Perspectives 5*(1), 159–170.

Sacerdote, B. (2001). Peer effects with random assignment: Results for Dartmouth roommates. *The Quarterly Journal of Economics 116*(2), 681–704.

Smith, J., J. Goodman, and M. Hurwitz (2020). The economic impact of access to public four-year colleges. National Bureau of Economic Research.

Statistics Canada (2017). Variant of classification of instructional programs 2016 - STEM and BHASE (non-STEM) groupings. https://www150.statcan.gc.ca/n1/daily-quotidien/171018/dq171018g-eng.htm. Accessed: June 2020.

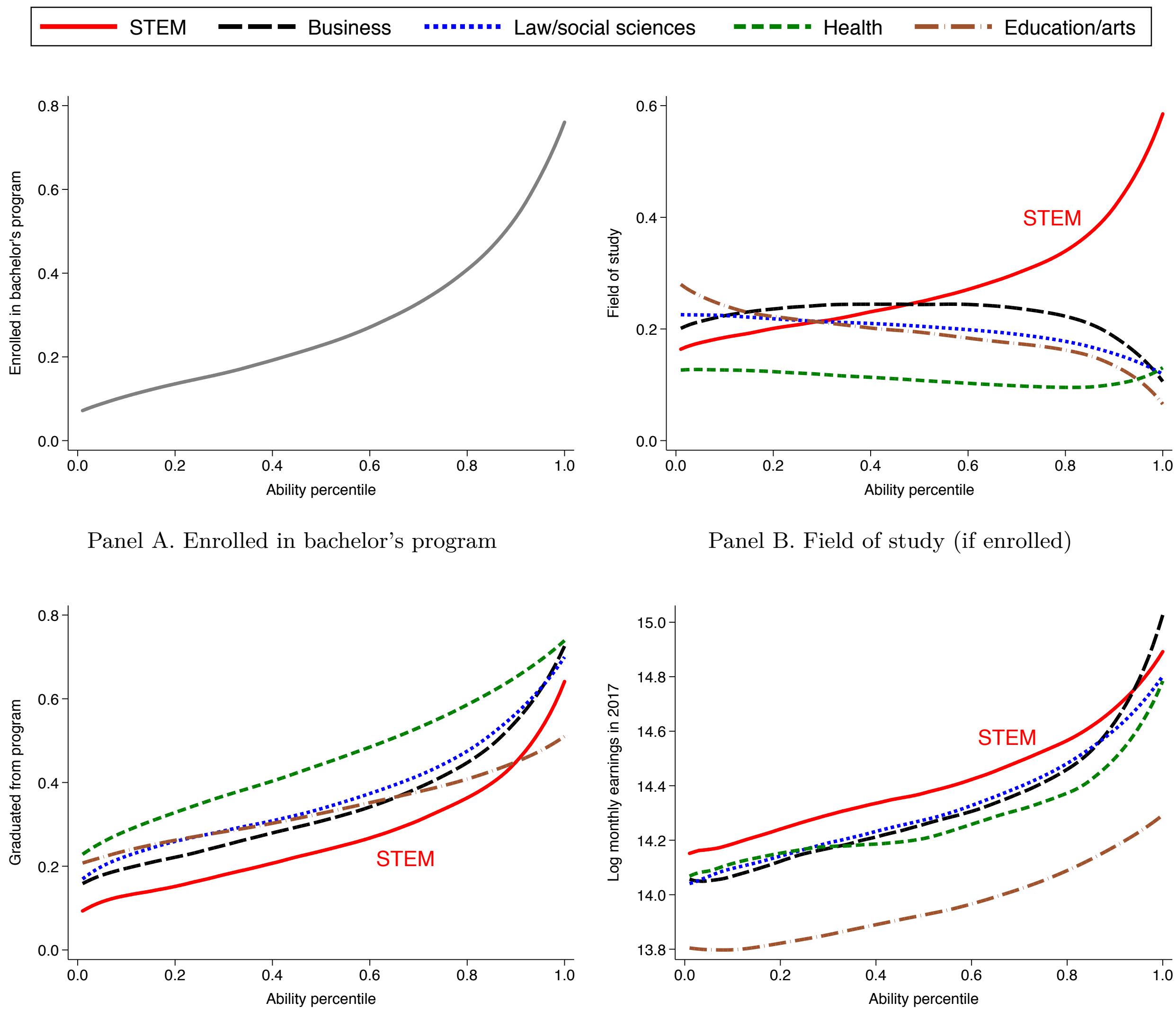
Stinebrickner, R. and T. R. Stinebrickner (2014). A major in science? initial beliefs and final outcomes for college major and dropout. *Review of Economic Studies 81*(1), 426–472.

Webber, D. A. (2014). The lifetime earnings premia of different majors: Correcting for selection based on cognitive, noncognitive, and unobserved factors. *Labour Economics 28*, 14–23.

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

Zimmerman, S. D. (2019). Elite colleges and upward mobility to top jobs and top incomes. *American Economic Review 109*(1), 1–47.

Figures



Panel A. Enrolled in bachelor's program

Panel B. Field of study (if enrolled)

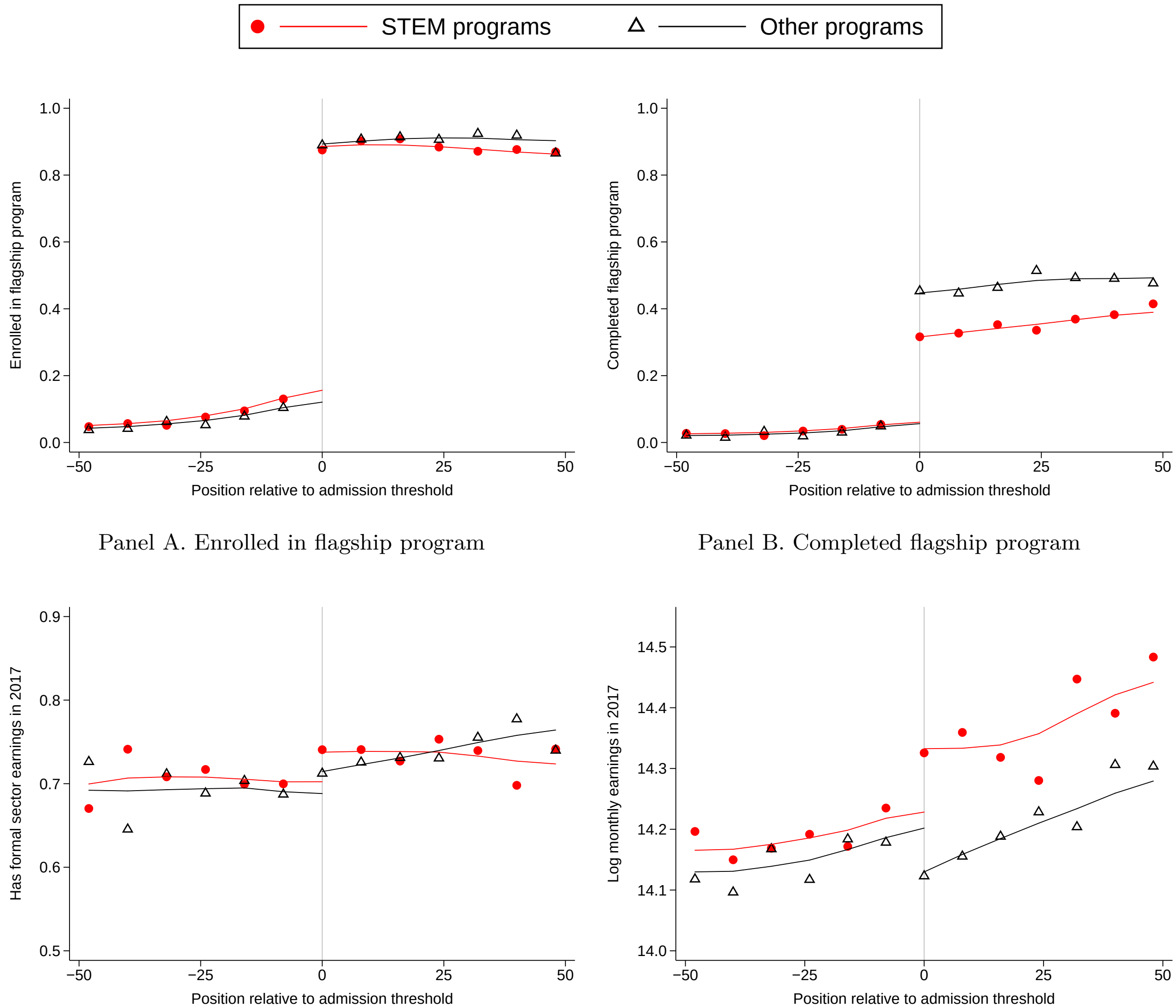
Panel C. Graduated from program (if enrolled)

Panel D. Log monthly earnings in 2017 (if graduated)

Figure 1. College and earnings outcomes by ability and degree field

Notes: This figure displays college and earnings outcomes by ability and degree field using Colombian administrative data. The sample in Panel A includes all students who took the ICFES exam in 1998–2004. Panels B–C include the subset of exam takers who enrolled in bachelor's program in our Ministry of Education data. Panel D includes the subset of exam takers who completed a bachelor's degree.

The x-axis in each panel is an applicant's graduation-anchored ability percentile (see Section 1.3). The variable on the y-axis is listed in the panel title. Lines are predicted values from local linear regressions of the outcome variables on ability percentile. In Panels B–D, lines depict the degree field groups listed in the legend, which we define using the Ministry of Education's categorization of programs into nine areas. STEM includes the Ministry's engineering and natural science areas. Education/arts includes the education, fine arts, agronomy, and religion areas.

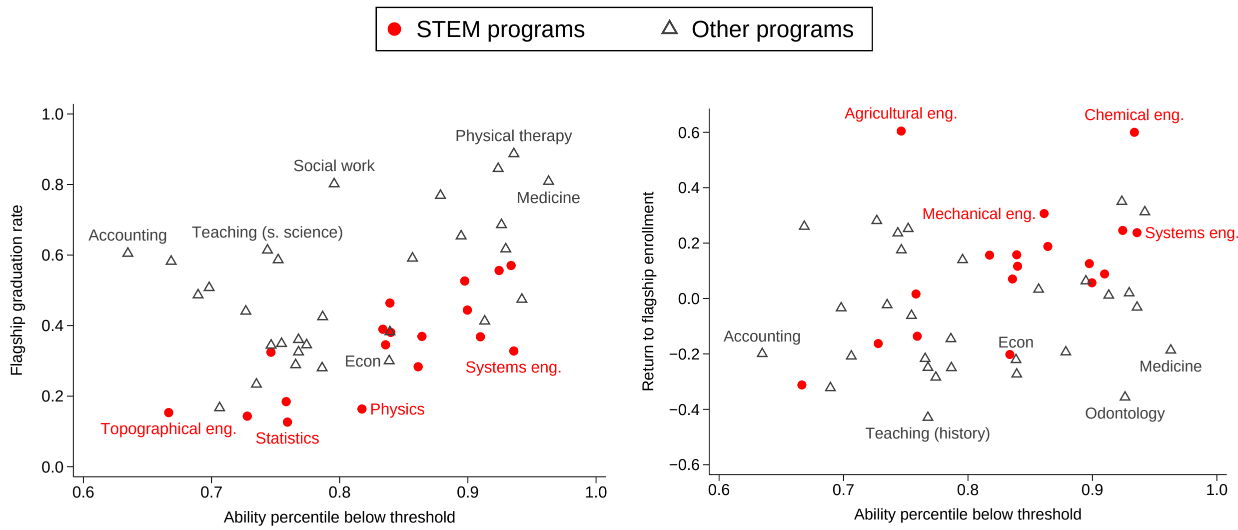


Panel C. Employed in formal sector in 2017

Panel D. Log monthly earnings in 2017

Figure 2. Returns to flagship admission

Notes: This figure presents RD graphs of the effects of admission to flagship programs. The x-axis in each panel is an applicant's position in their application pool normalized to zero at the threshold. The variable on the y-axis is listed in the panel title. Markers depict means of the dependent variable in bins of eight positions. Lines are predicted values from local linear regressions estimated separately above and below the threshold with a triangular kernel and a 30-rank bandwidth. Red circles and lines show estimates for STEM applicants. Black triangles and lines show estimates for applicants to other programs.

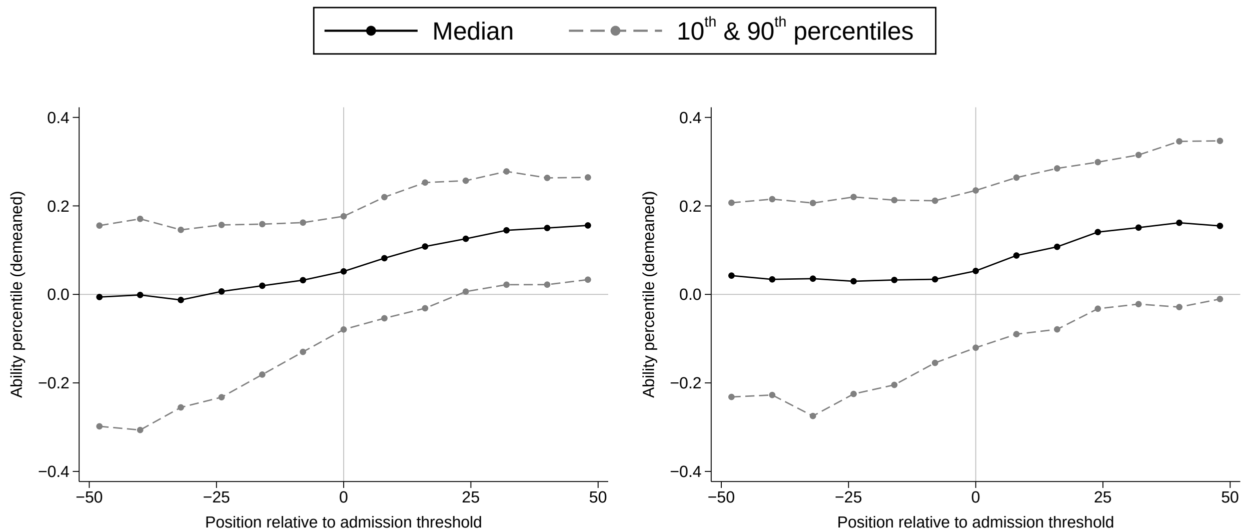


Panel A. Flagship graduation rate

Panel B. Return to flagship enrollment

FIGURE 3. Graduation and earnings returns to enrollment by program selectivity

Notes: This figure plots graduation rates and earnings returns for marginal enrollees in each of the 48 programs in our sample. The x-axis is the graduation-anchored ability percentile for applicants 1–5 positions below the admission thresholds in each program. The y-axis is the program's graduation rate (Panel A) or earnings returns to enrollment (Panel B), which are β coefficients from separate estimations of equation (3) for each program. The red circles are STEM programs, and the hollow triangles are other programs.

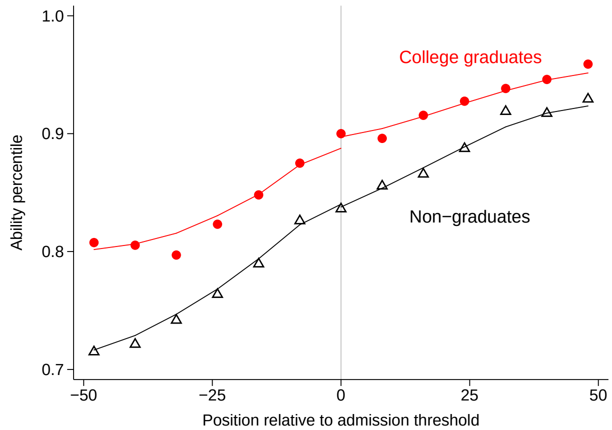


Panel A. STEM applicants

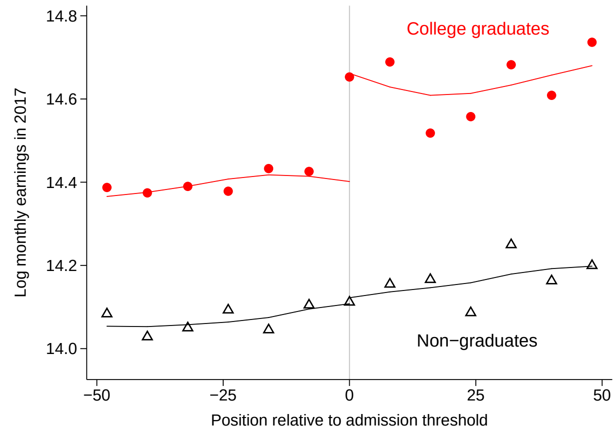
Panel B. Other applicants

FIGURE 4. Distribution of ability within application pools

Notes: This figure shows variation in ability conditional on admission rank. The x-axis in each panel is an applicant's position in their application pool normalized to zero at the threshold. The y-axis is an applicant's graduation-anchored ability percentile (see Section 1.3), which we demean for each application pool. The solid line shows the median ability percentile within 8-rank bins of the admission score. Dashed lines show the 10^{th} and 90^{th} percentiles of ability in these bins. Panel A includes STEM applicants, and Panel B includes applicants to other programs.



Panel A. Ability percentile



Panel B. Log monthly earnings in 2017

FIGURE 5. STEM returns by college degree attainment

Notes: This figure plots ability and earnings by college degree attainment. The x-axis in each panel is an applicant's position in their application pool normalized to zero at the threshold. The y-axis is an individual's graduation-anchored ability percentile (Panel A) or log monthly earnings in 2017 (Panel B).

The sample includes all STEM applicants. Markers show the means of each variable in 8-rank bins of the admission score. Red circles include applicants who earned a college degree, and hollow triangles include those who did not. Lines are local linear regressions estimated separately above and below the thresholds for each sample.

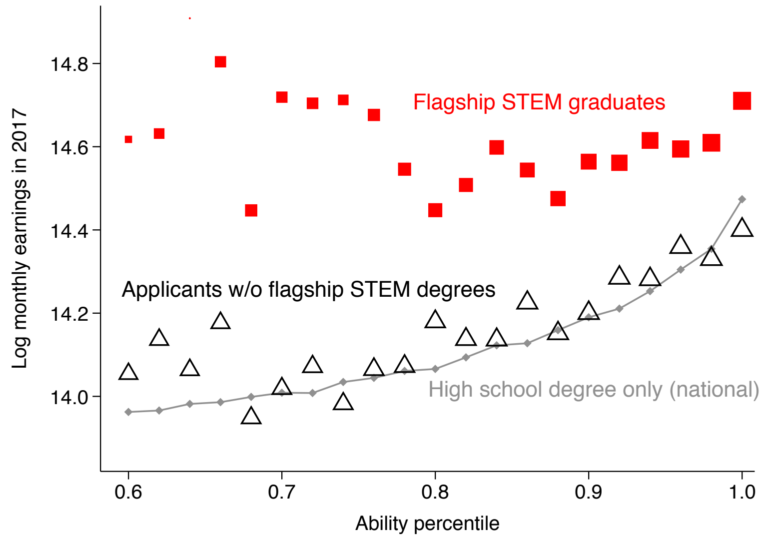


FIGURE 6. Log earnings by ability and educational attainment

Notes: This figure plots earnings by ability for three groups defined by educational attainment. The x-axis is an individual's graduation-anchored ability percentile. The y-axis is log monthly earnings in 2017.

Markers show mean earnings in two-percentile bins of ability. Red squares depict flagship STEM applicants who were above the admission threshold and completed the program. Hollow triangles depict STEM applicants who did not earn a degree from that program. For these two samples, marker sizes are proportional to the log number of observations. Small grey diamonds (connected with a line) depict all ICFES exam takers in our administrative data who did not earn a degree from any college. To compute mean earnings in this sample, we weight observations so that the distribution of ICFES exam years matches that for flagship STEM applicants.

Tables

Table 1. Relationship between log earnings and within-program ability percentile

	(A)	(B)	(C)	(D)
	All programs		Programs at flagship universities	
Degree field	Avg monthly earnings (2017 USD)	Earnings/ ability gradient	Avg monthly earnings (2017 USD)	Earnings/ ability gradient
All degrees	872	0.180	821	0.123
		(0.006)		(0.010)
STEM	1,023	0.145	983	0.072
		(0.011)		(0.016)
Business	877	0.267	709	0.244
		(0.013)		(0.020)
Law/social sciences	869	0.231	793	0.195
		(0.014)		(0.028)
Health	762	0.123	848	0.103
		(0.014)		(0.024)
Education/arts	525	0.109	491	0.073
		(0.017)		(0.024)
N (all degrees)	208,049	208,049	74,711	74,711

Notes: This table shows how earnings vary with ability among individuals who graduated from the same college program in the same year. The sample in columns (A)–(B) includes students who took the ICFES exam in 1998–2004 and graduated from a bachelor's program in our Ministry of Education data. We exclude observations from program/graduation-year pairs with fewer than ten graduates. Columns (C)–(D) include the subset of these individuals who graduated from a flagship university. We define flagships as the public university in each region with the highest mean graduation-anchored ability percentile. We include any flagship programs at campuses in other regions.

Columns (A) and (C) show average monthly earnings converted to 2017 U.S. dollars. Columns (B) and (D) display the coefficients from a regression of log monthly earnings in 2017 on graduation-anchored ability percentile. We renormalize this ability measure to reflect an individual's percentile in the distribution of students who graduated from the same college program in the same year. The first row presents estimates that pool across all degree fields. The bottom five rows show estimates for the degree field groups defined in Figure 1.

Parentheses contain standard errors clustered at the individual level.

Table 2. Programs and applicants in sample

	(A)	(B)	(C)	(D)	(E)	(F)	(G)	(H)
							Among admits	
Group	Faculty area	Program examples	# progs	# prog-cohorts	Total applied	Total admits	Yield rate	Ability pctile
STEM	Engineering	Civil, Mechanical, Systems	12	74	10,418	4,057	0.88	0.90
	Natural sciences	Chemistry, Physics, Math	6	33	5,604	1,603	0.89	0.90
Other	Administration	Accounting, Business	5	21	3,797	1,934	0.92	0.78
	Health	Dentistry, Medicine, Nursing	7	35	9,385	1,399	0.90	0.95
	Humanities	History, Teaching licenses	10	31	3,990	1,306	0.92	0.84
	Integrated arts	Architecture, Visual Arts	5	30	3,059	1,135	0.90	0.83
	Social sciences	Economics, Sociology	3	18	3,208	993	0.89	0.89
Total			48	242	39,461	12,427	0.90	0.87

Notes: This table displays summary statistics on the programs and applicants in our sample. Column (A) groups programs by their faculty area at Univalle, and column (B) lists examples of programs in each area. Columns (C)–(D) list the number of programs and program/application cohort pairs. See Appendix Table A1 for a full list of programs.

Columns (E)–(F) display the total number of applicants and admitted students in our sample. Columns (G)–(H) show the proportion enrolled (yield rate) and mean graduation-anchored ability percentile among admits.

TABLE 3. Returns to STEM and other flagship programs

	(A)	(B)	(C)	(D)
	STEM programs		Other programs	
Dependent variable	Mean below threshold	RD coef	Mean below threshold	RD coef
Ability percentile	0.848	0.004	0.815	0.007
		(0.005)		(0.006)
Panel A. Reduced form regressions				
Enrolled in flagship program	0.149	0.746	0.106	0.784
		(0.015)		(0.013)
Completed flagship program	0.061	0.256	0.053	0.390
		(0.016)		(0.015)
Employed in formal sector in 2017	0.696	0.030	0.686	0.028
		(0.020)		(0.020)
Log monthly earnings in 2017	14.232	0.100	14.188	−0.037
		(0.046)		(0.041)
Panel B. 2SLS regressions				
Flagship graduation rate		0.344		0.498
		(0.020)		(0.018)
Return to flagship enrollment		0.133		−0.047
		(0.060)		(0.050)
N	657	6,699	813	7,664

Notes: This table displays RD coefficients from separate regressions for STEM and other applicants. Column (A) presents means of the dependent variables listed in the row header for STEM applicants 1–5 positions below the admission thresholds. Column (B) presents RD coefficients using a sample of STEM applicants within 30 positions of the admission thresholds. Columns (C)–(D) are defined similarly using applicants to other programs.

The dependent variable in the first row is an applicant's graduation-anchored ability percentile (see Section 1.3). Panel A displays reduced-form RD coefficients, θ, from equation (2) using the dependent variable listed in the row header. Panel B displays 2SLS RD coefficients, β, from equation (3) using flagship program completion and log monthly earnings in 2017 as dependent variables.

Parentheses contain standard errors clustered at the individual level.

TABLE 4. Heterogeneity in flagship returns by ability

	(A)	(B)	(C)	(D)	(E)	(F)
	STEM programs			Other programs		
Dependent variable	Lower ability	High ability	p value diff	Lower ability	High ability	p value diff
Ability percentile below threshold	0.767	0.917		0.727	0.907	
Panel A. Reduced form regressions						
Enrolled in flagship program	0.761	0.747	0.637	0.815	0.759	0.036
	(0.021)	(0.021)		(0.018)	(0.020)	
Completed flagship program	0.203	0.297	0.004	0.370	0.399	0.355
	(0.022)	(0.024)		(0.022)	(0.023)	
Employed in formal sector in 2017	0.008	0.046	0.363	0.030	0.023	0.853
	(0.031)	(0.029)		(0.029)	(0.028)	
Log monthly earnings in 2017	0.196	0.042	0.101	−0.041	−0.009	0.714
	(0.069)	(0.065)		(0.061)	(0.060)	
Panel B. 2SLS regressions						
Flagship graduation rate	0.267	0.397	0.001	0.454	0.526	0.044
	(0.026)	(0.029)		(0.024)	(0.027)	
Return to flagship enrollment	0.257	0.057	0.096	−0.048	−0.013	0.729
	(0.088)	(0.084)		(0.070)	(0.076)	
N	3,390	3,309		3,886	3,778	

Notes: This table displays RD coefficients from separate regressions for lower- and high-ability applicants. In columns (A)–(C), we define lower- and high-ability STEM applicants from a regression of graduation-anchored ability on admission rank and application pool dummies. The sample in column (A) includes applicants whose residuals from this regression are below the median in their application pool. Column (B) includes applicants with above-median residuals from this regression. Column (C) displays the p value from an F test that the RD coefficients are equal for the two ability groups. Columns (D)–(F) are defined similarly using applicants to other programs.

The first row shows the mean graduation-anchored ability percentile for applicants 1–5 positions below the admission thresholds. Panel A displays reduced-form RD coefficients, θ, from equation (2) using the dependent variable listed in the row header. Panel B displays 2SLS RD coefficients, β, from equation (3) using flagship program completion and log monthly earnings in 2017 as dependent variables. All regressions include only applicants within 30 positions of the admission thresholds.

Parentheses contain standard errors clustered at the individual level.

Table 5. Number of students admitted to STEM programs

Quota variation	Program	Aug 1999	Jan 2000	Aug 2000	Jan 2001	Aug 2001	Jan 2002	Aug 2002	Jan 2003	Aug 2003	Jan 2004
		Number admitted by semester of application									
Program expansions	Biology	**101**		**99**		**82**	43	**92**	45	53	62
	Systems Eng.	62		82		**126**		61		63	
Tracking admissions	Chemical Eng.	61		**130**		66		43	41	39	36
	Electrical Eng.	56		**127**		57		45	51	49	45
	Electronic Eng.	64		**141**		71		54	44	55	46
	Mechanical Eng.	62		67		**123**		56	50	44	42
Minimal	Other programs (mean)	60	45	62	44	62	50	63	47	63	46

Notes: This table shows the number of students in our sample who were admitted to the flagship's STEM programs. Columns denote the semester of application, which we observe from August 1999 to January 2004.

The first six rows show programs in which the admission quotas changed significantly during this period. This includes two programs with class size expansions, and four programs that used "tracking" admissions (see Section 5.2). The last row shows the mean number of admits for the other 12 STEM programs in our sample.

Bold numbers are cohorts that we define as having large quotas for our binary measure of L_{mt} (see Section 5.3).

Table 6. Effects of quota variation on flagship STEM returns

	(A)	(B)	(C)	(D)	(E)
	RD estimates			RD diff-in-diff coefs	
Dependent variable	Large quota cohorts	Small quota cohorts	p value diff	Effect of large quota	Effect of 60 extra admits
Ability percentile below threshold	0.833	0.918		−0.082	−0.086
				(0.026)	(0.025)
Panel A. Reduced form regressions					
Enrolled in flagship program	0.779	0.613	0.003	0.153	0.149
	(0.041)	(0.036)		(0.059)	(0.061)
Completed flagship program	0.266	0.274	0.928	−0.022	−0.015
	(0.078)	(0.044)		(0.063)	(0.062)
Employed in formal sector in 2017	0.046	0.025	0.778	0.012	0.016
	(0.061)	(0.043)		(0.088)	(0.065)
Log monthly earnings in 2017	0.413	0.067	0.007	0.307	0.176
	(0.072)	(0.100)		(0.141)	(0.130)
Panel B. 2SLS regressions					
Flagship graduation rate	0.342	0.446	0.354	−0.125	−0.116
	(0.094)	(0.064)		(0.072)	(0.057)
Return to flagship enrollment	0.546	0.114	0.023	0.401	0.211
	(0.104)	(0.158)		(0.216)	(0.206)
N	671	1,927		6,699	6,699

Notes: This table shows how flagship STEM returns changed with the size of the admission quotas.

Columns (A)–(C) present RD coefficients from separate regressions for cohorts with large and small quotas. The sample for column (A) includes the cohorts in bold for the six STEM programs listed in Table 5. The sample for column (B) includes the non-bold cohorts of these six programs. We weight observations in these regressions so that the sample in each column has the same distribution of programs as in the combined sample for both columns. Column (C) displays the p value from an F test that the RD coefficients are equal for the two cohort groups.

Columns (D)–(E) present estimates from our RD differences-in-differences specification. The sample includes applicants to all STEM programs. In column (D), the variable of interest, L_{mt}, is an indicator for the cohorts in bold in Table 5. In column (E), we define L_{mt} as the total number of admits in each program/cohort divided by 60.

In the first row, columns (A)–(B) show the mean graduation-anchored ability percentile for applicants 1–5 positions below the admission thresholds, and columns (D)–(E) show estimates of π from equation (5) with mean ability below the threshold as the dependent variable. Panel A displays reduced-form estimates of θ from equation (2) (columns (A)–(C)) and of π from equations (4)–(5) (columns (D)–(E)) using the dependent variable in the row header. Panel B displays 2SLS estimates of β from equation (3) (columns (A)–(C)) and of the RD differences-in-differences coefficient (columns (D)–(E)) using flagship program completion and log monthly earnings in 2017 as dependent variables. All regressions include only applicants within 30 positions of the admission thresholds.

Parentheses contain standard errors clustered at the program/cohort level.

TABLE 7. Effects of flagship STEM admission on college choices and graduation outcomes

	(A)	(B)	(C)	(D)	(E)
	Lower-ability applicants		High-ability applicants		
Dependent variable	Mean below threshold	RD coef	Mean below threshold	RD coef	p value diff
Enrolled in any college program	0.769	0.170 (0.024)	0.895	0.084 (0.017)	0.004
Enrolled in any STEM program	0.475	0.428 (0.028)	0.616	0.331 (0.026)	0.011
Completed any college program	0.363	0.081 (0.033)	0.489	0.096 (0.032)	0.757
Completed any STEM program	0.158	0.158 (0.027)	0.266	0.198 (0.029)	0.323
N	303	3,390	354	3,309	

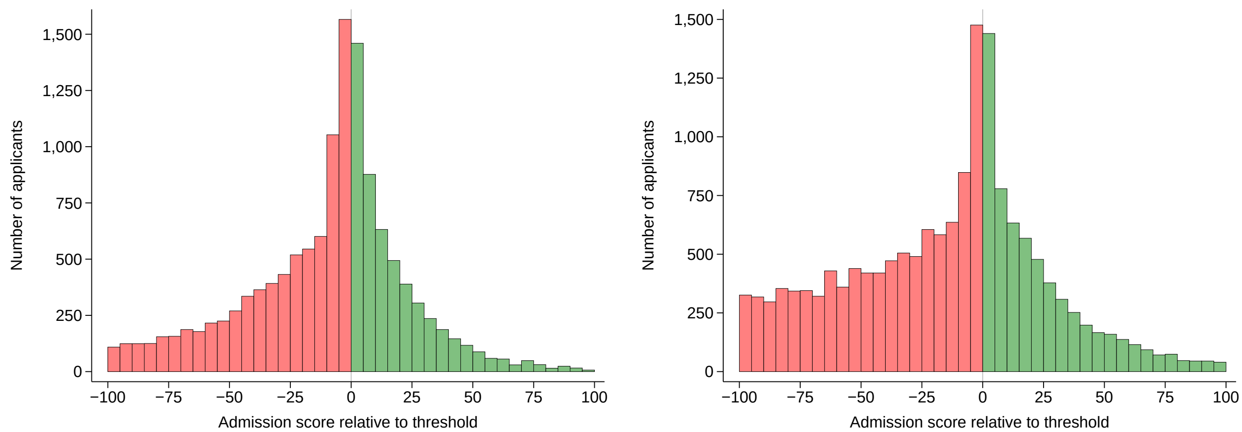
Notes: This table displays RD coefficients from separate regressions for lower- and high-ability applicants, which we define as in Table 4. Columns (A)–(B) include lower-ability STEM applicants, and columns (C)–(D) include high-ability STEM applicants. All regressions include only applicants within 30 positions of the admission thresholds.

Columns (A) and (C) display means of the dependent variables listed in the row header for applicants 1–5 positions below the admission thresholds. We define STEM majors as bachelor's programs in engineering and natural sciences using the Ministry of Education's classification of program areas. Columns (B) and (D) display reduced-form RD coefficients, θ, from equation (2). Column (E) displays the p value from an F test that the RD coefficients are equal for the two ability groups.

Parentheses contain standard errors clustered at the individual level.

Appendix — For Online Publication

A. Appendix figures and tables



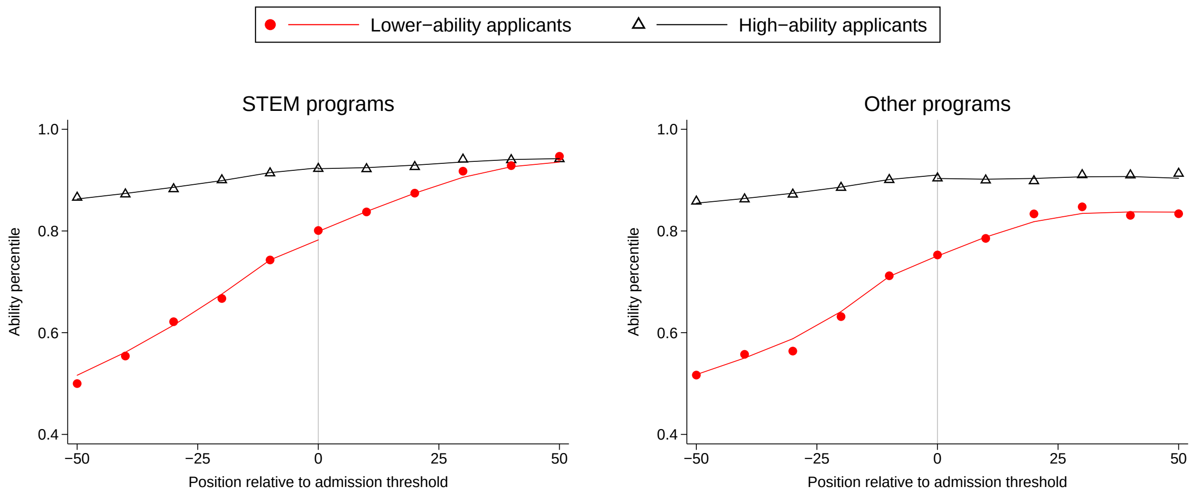
Panel A. STEM applicants

Panel B. Other applicants

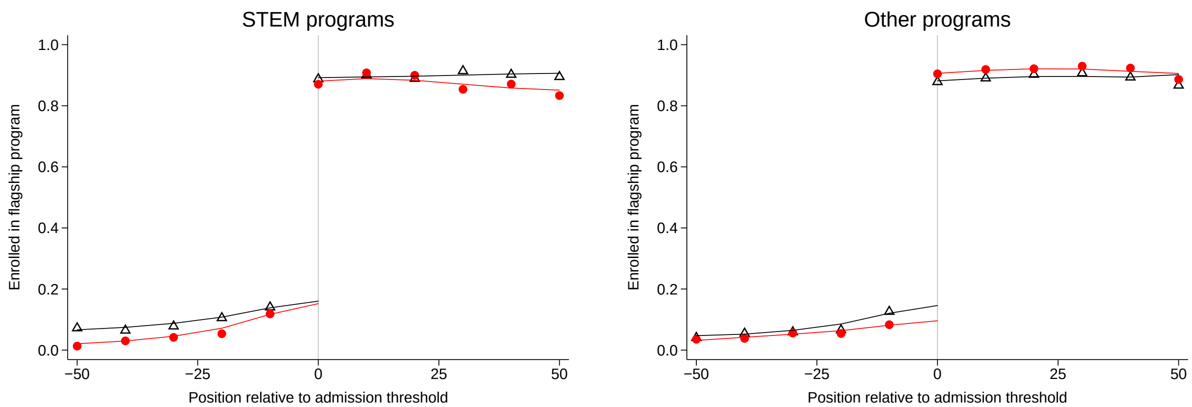
Figure A1. Density of admission scores relative to the threshold

Notes: This figure shows the density of admission scores relative to the admission thresholds. The x-axis is a student's admission score normalized to zero at the threshold. The y-axis shows the number of applicants within five unit bins of the admission score. The graphs are limited to those with normalized scores between -100 and 100.

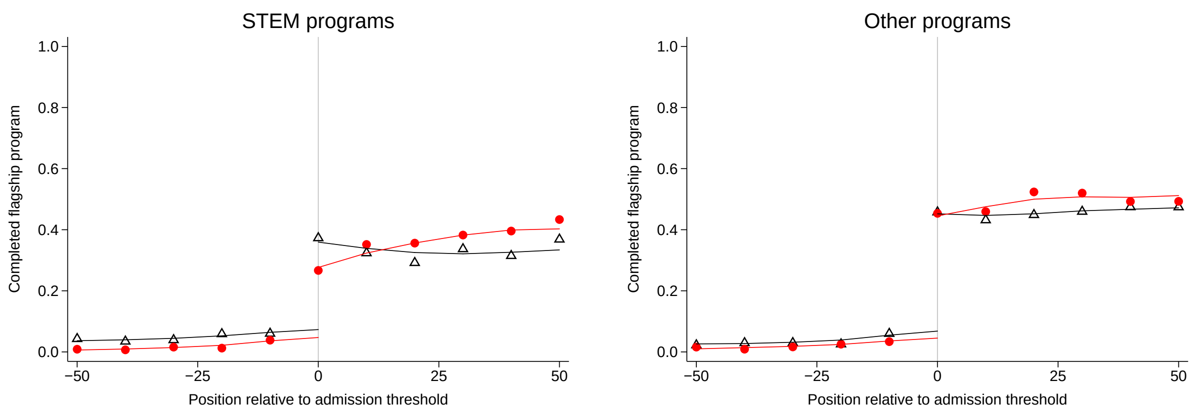
Panel A shows the distribution of admission scores for applicants to flagship STEM programs. Using the McCrary (2008) density test, the estimated discontinuity—i.e., the log difference in height at the threshold—is -0.049 with a standard error of 0.030. Panel B shows the distribution of admission scores for applicants to other flagship programs. The estimated density discontinuity is 0.017 with a standard error of 0.026.



Panel A. Ability percentile



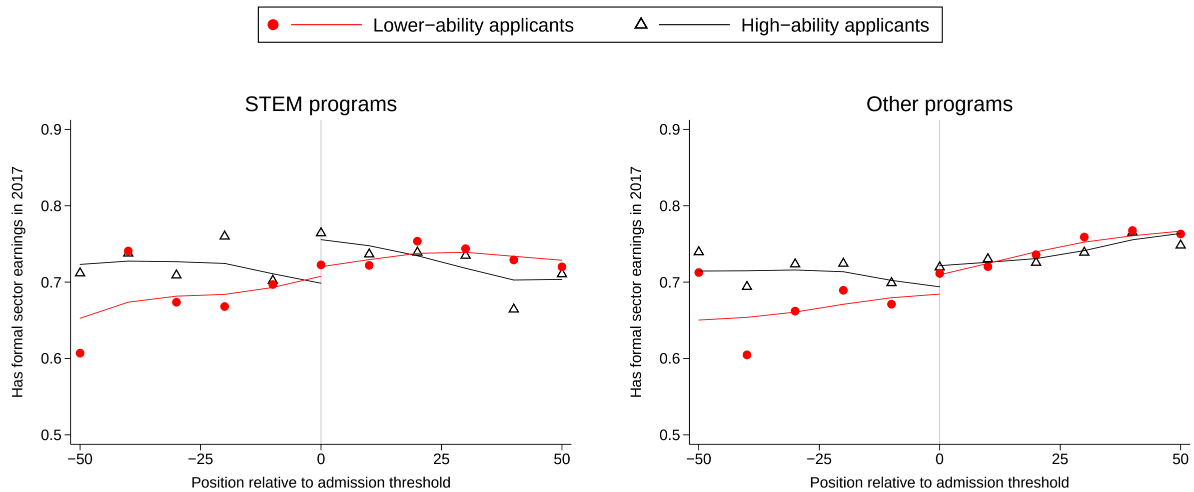
Panel B. Enrolled in flagship program



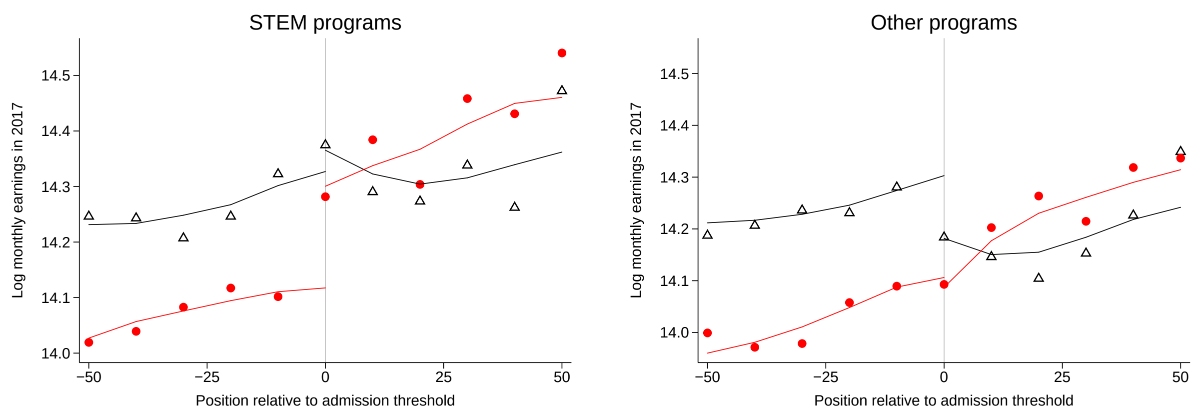
Panel C. Completed flagship program

FIGURE A2. Heterogeneity in returns to flagship programs by ability

Notes: This figure is continued on the next page.



Panel D. Employed in formal sector in 2017

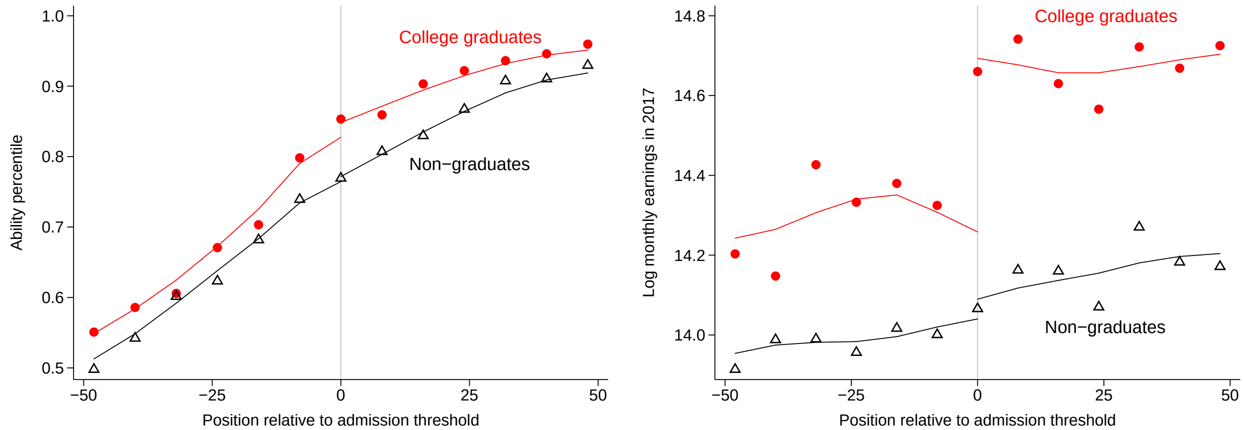


Panel E. Log monthly earnings in 2017

FIGURE A2. Heterogeneity in returns to flagship programs by ability (continued)

Notes: This figure presents RD graphs of the effects of admission to flagship programs. The x-axis in each panel is an applicant's position in their application pool normalized to zero at the threshold. The variable on the y-axis is listed in the panel title. Markers show the means of each variable in 10-rank bins of the admission score. Lines are predicted values from local linear regressions estimated separately above and below the threshold with a triangular kernel and a 30-rank bandwidth.

We display variable means and predicted values separately for high- and lower-ability applicants, defined as in Table 4. Black triangles and lines triangles show estimates for high-ability applicants. Red circles and lines show estimates for lower-ability applicants. Graphs in the left column include STEM applicants, and graphs in the right column include other applicants.



Panel A. Ability percentile   Panel B. Log monthly earnings in 2017

Figure A3. STEM returns by college degree attainment — Lower-ability applicants

Notes: This figure is identical to Figure 5, but we restrict the sample to lower-ability applicants, defined as in Table 4. This figure plots ability and earnings by college degree attainment. The x-axis in each panel is an applicant's position in their application pool normalized to zero at the threshold. The y-axis is an individual's graduation-anchored ability percentile (Panel A) or log monthly earnings in 2017 (Panel B).

The sample includes lower-ability applicants to STEM programs, as in column (A) of Table 4. Markers show the means of each variable in 8-rank bins of the admission score. Red circles include applicants who earned a college degree, and hollow triangles include those who did not. Lines are local linear regressions estimated separately above and below the thresholds for each sample.

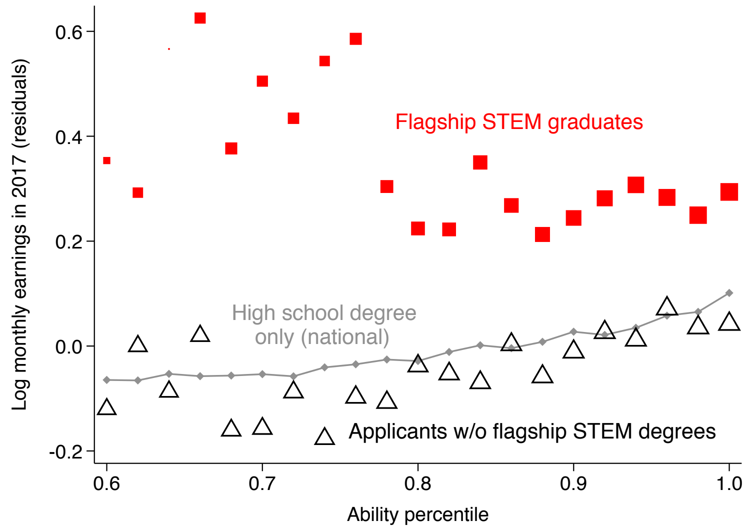


FIGURE A4. Log earnings residuals by ability and educational attainment

Notes: This figure is similar to Figure 6, but the dependent variable is earnings residuals rather than raw earnings. The x-axis is an individual's graduation-anchored ability percentile. The y-axis depicts residuals from a regression of log monthly earnings in 2017 on a vector of individual covariates. These covariates include gender, age, and dummies for high schools, mother's education categories, father's education categories, family income bins, and ICFES exam years. We estimate this regression using the full sample for each dataset.

Markers show mean earnings in two-percentile bins of ability. Red squares depict flagship STEM applicants who were above the admission threshold and completed the program. Hollow triangles depict STEM applicants who did not earn a degree from that program. For these two samples, marker sizes are proportional to the log number of observations. Small grey diamonds (connected with a line) depict all ICFES exam takers in our administrative data who did not earn a degree from any college. To compute mean earnings in this sample, we weight observations so that the distribution of ICFES exam years matches that for flagship STEM applicants.

Table A1. Programs in sample

	(A)		(B)	(C)	(D)	(E)
Group	Faculty area	#	Program	Application cohorts	Total applied	Main RD sample
STEM	Engineering	1	Agricultural Engineering	6	532	313
		2	Chemical Engineering	7	1,220	478
		3	Civil Engineering	7	590	405
		4	Electrical Engineering	7	717	380
		5	Electronic Engineering	7	1,027	407
		6	Industrial Engineering	7	1,183	423
		7	Materials Engineering	6	857	379
		8	Mechanical Engineering	7	849	443
		9	Sanitary Engineering	4	541	274
		10	Statistics	5	627	254
		11	Systems Engineering	5	1,758	323
		12	Topographical Engineering	6	517	306
	Natural sciences	13	Biology	8	2,021	567
		14	Chemical Technology (day)	3	883	238
		15	Chemical Technology (night)	3	295	212
		16	Chemistry	7	1,073	473
		17	Math	3	481	259
		18	Physics	9	851	565
Other	Administration	19	Accounting (day)	5	845	250
		20	Accounting (night)	4	758	274
		21	Business (day)	5	1,065	275
		22	Business (night)	5	770	299
		23	Foreign Trade	2	359	107
	Health	24	Audiology	5	579	294
		25	Bacteriology	5	1,657	301
		26	Dentistry	5	818	286
		27	Medicine	5	2,551	327
		28	Nursing	5	1,149	261
		29	Occupational Therapy	5	889	286
		30	Physical Therapy	5	1,742	297
	Humanities	31	History	4	531	190
		32	Recreation	2	228	123
		33	Social Work	4	1,016	233
		34	Teaching (Elem. S. Science)	2	154	108
		35	Teaching (Foreign Language, day)	2	188	114
		36	Teaching (Foreign Language, night)	2	107	93
		37	Teaching (History)	4	596	213
		38	Teaching (Literature)	4	588	260
		39	Teaching (Philosophy)	4	411	261
		40	Teaching (Social Science)	3	171	139
	Integrated arts	41	Architecture	6	1,346	311
		42	Communication	5	356	268
		43	Dramatic Arts	9	363	336
		44	Teaching (Music)	5	571	332
		45	Visual Arts	5	423	280
	Social sciences	46	Economics	9	983	585
		47	Psychology	5	1,264	323
		48	Sociology	4	961	238
Total				242	39,461	14,363

Notes: Columns (A)–(B) list each Univalle program in our sample and its faculty area (see Section 2.2). Column (C) shows the total number of application cohorts from August 1999 to January 2004. Column (D) reports the total number of applicants in our sample, and column (E) shows the number of applicants within 30 positions of the admission thresholds.

Table A2. Graduation-anchored ICFES subject scores

	(A)	(B)
	Dependent variable: Graduated from college	
Subject	1998–1999 ICFES exam takers	2000–2004 ICFES exam takers
Language	0.014	0.025
	(0.002)	(0.001)
Math knowledge	0.030	
	(0.001)	
Math aptitude	−0.007	
	(0.001)	
Math		0.008
		(0.001)
Biology	−0.003	0.021
	(0.002)	(0.001)
Chemistry	0.057	0.047
	(0.002)	(0.001)
Physics	0.019	0.007
	(0.002)	(0.001)
Social science	0.013	
	(0.002)	
History		0.020
		(0.001)
Geography		0.014
		(0.001)
Philosophy		0.024
		(0.001)
Interdisciplinary		0.008
		(0.001)
N	280,894	799,761

Notes: This table reports coefficients on ICFES subject scores from the regressions that we use to compute our graduation-anchored measure of ability (see Section 1.3). The sample includes all Colombian high school graduates who took the ICFES exam in 1998–2004 and enrolled in any university in our administrative records. We exclude individuals who applied to Univalle during our sample period.

The dependent variable is an indicator equal to one for individuals who graduated from their college program. Column (A) includes students who took the ICFES exam in 1998–1999. Column (B) includes students who took the ICFES exam in 2000–2004. The exam subjects changed between the two versions as shown by the reported coefficients. We normalize each subject score to have standard deviation one within each sample.

Parentheses contain standard errors clustered at the individual level.

Table A3. Flagship returns by RD bandwidth — STEM programs

	(A)	(B)	(C)	(D)	(E)
		RD coefficients by bandwidth, h (in admission ranks)			
	Mean below threshold	$h = 45$	$h = 30$	$h = 15$	CCT
Ability percentile	0.848	0.012	0.004	−0.001	0.002
		(0.005)	(0.005)	(0.007)	(0.006)
Panel A. Reduced form regressions					
Enrolled in flagship program	0.149	0.766	0.746	0.716	0.725
		(0.012)	(0.015)	(0.021)	(0.018)
Completed flagship program	0.061	0.259	0.256	0.242	0.256
		(0.014)	(0.016)	(0.022)	(0.016)
Employed in formal sector in 2017	0.696	0.030	0.030	0.033	0.035
		(0.017)	(0.020)	(0.028)	(0.020)
Log monthly earnings in 2017	14.232	0.078	0.100	0.053	0.092
		(0.038)	(0.046)	(0.063)	(0.047)
Panel B. 2SLS regressions					
Flagship graduation rate		0.338	0.344	0.338	0.344
		(0.017)	(0.020)	(0.028)	(0.020)
Return to flagship enrollment		0.101	0.133	0.074	0.123
		(0.048)	(0.060)	(0.085)	(0.063)
N	657	8,994	6,699	3,789	5,215

Notes: This table displays RD coefficients for STEM applicants with different sample bandwidths. Column (A) presents means of the dependent variables listed in the row header for STEM applicants 1–5 positions below the admission thresholds. Columns (B)–(D) presents RD coefficients using a sample of STEM applicants within h positions of the admission thresholds, where h is listed in the column header. Column (E) uses the sample bandwidth from the benchmark method of Calonico et al. (2014) for each outcome.

The dependent variable in the first row is an applicant's graduation-anchored ability percentile. Panel A displays reduced-form RD coefficients, θ, from equation (2) using the dependent variable listed in the row header. Panel B displays 2SLS RD coefficients, β, from equation (3) using flagship program completion and log monthly earnings in 2017 as dependent variables.

Parentheses contain standard errors clustered at the individual level.

Table A4. Flagship returns by RD bandwidth — Other programs

	(A)	(B)	(C)	(D)	(E)
		RD coefficients by bandwidth, h (in admission ranks)			
	Mean below threshold	$h = 45$	$h = 30$	$h = 15$	CCT
Ability percentile	0.815	0.014	0.007	0.005	0.005
		(0.005)	(0.006)	(0.007)	(0.006)
Panel A. Reduced form regressions					
Enrolled in flagship program	0.106	0.788	0.784	0.765	0.781
		(0.011)	(0.013)	(0.018)	(0.013)
Completed flagship program	0.053	0.393	0.390	0.393	0.394
		(0.013)	(0.015)	(0.021)	(0.014)
Employed in formal sector in 2017	0.686	0.014	0.028	0.026	0.009
		(0.017)	(0.020)	(0.026)	(0.017)
Log monthly earnings in 2017	14.188	−0.055	−0.037	−0.007	−0.050
		(0.034)	(0.041)	(0.055)	(0.033)
Panel B. 2SLS regressions					
Flagship graduation rate		0.499	0.498	0.514	0.501
		(0.015)	(0.018)	(0.024)	(0.016)
Return to flagship enrollment		−0.070	−0.047	−0.009	−0.063
		(0.042)	(0.050)	(0.068)	(0.040)
N	813	10,026	7,664	4,439	6,659

Notes: This table displays RD coefficients for applicants to non-STEM programs with different sample bandwidths. Column (A) presents means of the dependent variables listed in the row header for non-STEM applicants 1–5 positions below the admission thresholds. Columns (B)–(D) presents RD coefficients using a sample of non-STEM applicants within h positions of the admission thresholds, where h is listed in the column header. Column (E) uses the sample bandwidth from the benchmark method of Calonico et al. (2014) for each outcome.

The dependent variable in the first row is an applicant's graduation-anchored ability percentile. Panel A displays reduced-form RD coefficients, θ, from equation (2) using the dependent variable listed in the row header. Panel B displays 2SLS RD coefficients, β, from equation (3) using flagship program completion and log monthly earnings in 2017 as dependent variables.

Parentheses contain standard errors clustered at the individual level.

TABLE A5. RD balance tests

	(A)	(B)	(C)	(D)
			RD coefficients	
Dependent variable	N	Mean below threshold	STEM applicants	Other applicants
Ability percentile	14,363	0.830	0.004	0.007
			(0.005)	(0.006)
Age	14,350	19.681	0.022	−0.155
			(0.114)	(0.120)
College educated father	12,847	0.423	−0.011	0.003
			(0.024)	(0.022)
College educated mother	13,690	0.357	0.021	0.003
			(0.022)	(0.020)
Family income > 2x min wage	13,820	0.591	0.016	0.035
			(0.023)	(0.021)
Female	14,363	0.450	−0.024	−0.004
			(0.020)	(0.020)
p value: Jointly zero			0.619	0.433

Notes: This table displays RD balance tests from separate regressions for STEM and other applicants. Column (A) shows the number of non-missing observations for applicants to all programs who were within 30 positions of the admission thresholds. Column (B) presents means of the dependent variables listed in the row header for applicants to all programs who were 1–5 positions below the admission thresholds. Column (B) presents RD coefficients using a sample of STEM applicants within 30 positions of the admission thresholds. Columns (C)–(D) display reduced-form RD coefficients, θ, from equation (2) using the dependent variable listed in the row header. Column (C) includes the applicants in column (A) who applied to STEM programs, and column (D) includes the applicants to other programs.

The last row reports p values from F tests that the coefficients on all covariates are jointly equal to zero.

Parentheses contain standard errors clustered at the individual level.

Table A6. Heterogeneity in flagship returns by ability (no graduation anchoring)

	(A)	(B)	(C)	(D)	(E)	(F)
	STEM programs			Other programs		
Dependent variable	Lower ability	High ability	p value diff	Lower ability	High ability	p value diff
Ability percentile below threshold	0.771	0.921		0.733	0.905	
Panel A. Reduced form regressions						
Enrolled in flagship program	0.751	0.746	0.877	0.810	0.767	0.109
	(0.022)	(0.021)		(0.018)	(0.020)	
Completed flagship program	0.225	0.274	0.139	0.371	0.396	0.432
	(0.022)	(0.024)		(0.022)	(0.023)	
Employed in formal sector in 2017	0.013	0.042	0.490	0.028	0.042	0.716
	(0.031)	(0.029)		(0.029)	(0.028)	
Log monthly earnings in 2017	0.165	0.064	0.276	−0.074	−0.003	0.404
	(0.064)	(0.066)		(0.063)	(0.058)	
Panel B. 2SLS regressions						
Flagship graduation rate	0.300	0.367	0.094	0.458	0.516	0.102
	(0.027)	(0.029)		(0.024)	(0.026)	
Return to flagship enrollment	0.216	0.087	0.279	−0.088	−0.005	0.414
	(0.081)	(0.087)		(0.072)	(0.074)	
N	3,390	3,309		3,886	3,778	

Notes: This table displays RD coefficients from separate regressions for lower- and high-ability applicants. This table is similar to Table 4, but we define ability based on the simple average of applicants' ICFES subject scores rather than the graduation-anchored average.

In columns (A)–(C), we define lower- and high-ability STEM applicants from a regression of mean ICFES percentile on admission rank and application pool dummies. The sample in column (A) includes applicants whose residuals from this regression are below the median in their application pool. Column (B) includes applicants with above-median residuals from this regression. Column (C) displays the p value from an F test that the RD coefficients are equal for the two ability groups. Columns (D)–(F) are defined similarly using applicants to other programs.

The first row shows the mean graduation-anchored ability percentile for applicants 1–5 positions below the admission thresholds. Panel A displays reduced-form RD coefficients, θ, from equation (2) using the dependent variable listed in the row header. Panel B displays 2SLS RD coefficients, β, from equation (3) using flagship program completion and log monthly earnings in 2017 as dependent variables. All regressions include only applicants within 30 positions of the admission thresholds.

Table A7. Heterogeneity in flagship returns by ability (post-2000 ICFES exam takers)

	(A)	(B)	(C)	(D)	(E)	(F)
	STEM programs			Other programs		
Dependent variable	Lower ability	High ability	p value diff	Lower ability	High ability	p value diff
Ability percentile below threshold	0.692	0.896		0.644	0.882	
Panel A. Reduced form regressions						
Enrolled in flagship program	0.779	0.731	0.179	0.848	0.776	0.017
	(0.025)	(0.026)		(0.019)	(0.023)	
Completed flagship program	0.188	0.292	0.008	0.355	0.398	0.255
	(0.026)	(0.029)		(0.026)	(0.027)	
Employed in formal sector in 2017	−0.014	0.015	0.560	0.027	0.031	0.946
	(0.037)	(0.035)		(0.034)	(0.034)	
Log monthly earnings in 2017	0.180	0.011	0.118	0.012	−0.010	0.828
	(0.078)	(0.075)		(0.070)	(0.072)	
Panel B. 2SLS regressions						
Flagship graduation rate	0.242	0.399	0.001	0.419	0.513	0.024
	(0.031)	(0.036)		(0.027)	(0.031)	
Return to flagship enrollment	0.229	0.016	0.127	0.014	−0.013	0.824
	(0.098)	(0.101)		(0.078)	(0.091)	
N	2,477	2,435		2,885	2,823	

Notes: This table displays RD coefficients from separate regressions for lower- and high-ability applicants. This table is similar to Table 4, but we restrict the sample to applicants who took the ICFES exam in 2000 or later.

In columns (A)–(C), we define lower- and high-ability STEM applicants from a regression of graduation-anchored ability on admission rank and application pool dummies. The sample in column (A) includes applicants whose residuals from this regression are below the median in their application pool. Column (B) includes applicants with above-median residuals from this regression. Column (C) displays the p value from an F test that the RD coefficients are equal for the two ability groups. Columns (D)–(F) are defined similarly using applicants to other programs.

The first row shows the mean graduation-anchored ability percentile for applicants 1–5 positions below the admission thresholds. Panel A displays reduced-form RD coefficients, θ, from equation (2) using the dependent variable listed in the row header. Panel B displays 2SLS RD coefficients, β, from equation (3) using flagship program completion and log monthly earnings in 2017 as dependent variables. All regressions include only applicants within 30 positions of the admission thresholds.

Parentheses contain standard errors clustered at the individual level.

Table A8. Effects of quota variation on flagship STEM returns (post-2000 ICFES exam takers)

	(A)	(B)	(C)	(D)	(E)
	RD estimates			RD diff-in-diff coefs	
Dependent variable	Large quota cohorts	Small quota cohorts	p value diff	Effect of large quota	Effect of 60 extra admits
Ability percentile below threshold	0.788	0.899		−0.074	−0.087
				(0.043)	(0.039)
Panel A. Reduced form regressions					
Enrolled in flagship program	0.809	0.604	0.001	0.199	0.212
	(0.019)	(0.053)		(0.065)	(0.070)
Completed flagship program	0.186	0.260	0.511	−0.057	−0.053
	(0.095)	(0.065)		(0.079)	(0.078)
Employed in formal sector in 2017	−0.026	0.040	0.469	−0.028	−0.023
	(0.074)	(0.059)		(0.105)	(0.082)
Log monthly earnings in 2017	0.374	0.007	0.026	0.226	0.121
	(0.094)	(0.130)		(0.171)	(0.147)
Panel B. 2SLS regressions					
Flagship graduation rate	0.230	0.430	0.163	−0.178	−0.198
	(0.106)	(0.097)		(0.088)	(0.070)
Return to flagship enrollment	0.476	0.012	0.048	0.265	0.114
	(0.113)	(0.205)		(0.251)	(0.230)
N	379	1,435		4,912	4,912

Notes: This table shows how flagship STEM returns changed with the size of the admission quotas. This table is similar to Table 6, but we restrict the sample to applicants who took the ICFES exam in 2000 or later.

Columns (A)–(C) present RD coefficients from separate regressions for cohorts with large and small quotas. The sample for column (A) includes the cohorts in bold for the six STEM programs listed in Table 5. The sample for column (B) includes the non-bold cohorts of these six programs. We weight observations in these regressions so that the sample in each column has the same distribution of programs as in the combined sample for both columns. Column (C) displays the p value from an F test that the RD coefficients are equal for the two cohort groups.

Columns (D)–(E) present estimates from our RD differences-in-differences specification. The sample includes applicants to all STEM programs. In column (D), the variable of interest, L_{mt}, is an indicator for the cohorts in bold in Table 5. In column (E), we define L_{mt} as the total number of admits in each program/cohort divided by 60.

In the first row, columns (A)–(B) show the mean graduation-anchored ability percentile for applicants 1–5 positions below the admission thresholds, and columns (D)–(E) show estimates of π from equation (5) with mean ability below the threshold as the dependent variable. Panel A displays reduced-form estimates of θ from equation (2) (columns (A)–(C)) and of π from equations (4)–(5) (columns (D)–(E)) using the dependent variable in the row header. Panel B displays 2SLS estimates of β from equation (3) (columns (A)–(C)) and of the RD differences-in-differences coefficient (columns (D)–(E)) using flagship program completion and log monthly earnings in 2017 as dependent variables. All regressions include only applicants within 30 positions of the admission thresholds.

Parentheses contain standard errors clustered at the program/cohort level.

Table A9. Differences-in-differences estimates of ability variation in flagship STEM returns

	(A)	(B)	(C)	(D)	(E)	(F)
	Dependent variable					
	Reduced form regressions				2SLS regressions	
	Enrolled in flagship	Finished flagship program	Formal employment	Log monthly earnings	Flagship grad rate	Return to flagship
Admitted	0.793	0.311	0.014	0.045	0.393	0.056
	(0.009)	(0.013)	(0.011)	(0.023)	(0.015)	(0.029)
Admitted × Ability	−0.003	0.004	−0.001	−0.003	0.006	−0.003
	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
N	16,022	16,022	16,022	11,302	16,022	11,302

Notes: This table presents estimates of ability variation in flagship STEM returns from a specification that exploits changes in the probability of admission across programs and cohorts. Specifically, the table display estimates of β and β^A from the following differences-in-differences specification:

$$Y_{imt} = \gamma_m X_i + \gamma_t X_i + \beta \text{Admitted}_{imt} + \beta^A (\text{Admitted}_{imt} \times \text{Ability}_i) + \epsilon_{imt}$$

The dependent variable, Y_{imt}, is an outcome for applicant i who applied to program m in application cohort, t. The dependent variables are the same as in Panels A–B of Table 6 and are listed in the column header. The variables of interest are an indicator for having an admission score above the threshold, Admitted_{imt}, and the interaction of this variable with the applicant's graduation-anchored ability percentile, $\text{Admitted}_{imt} \times \text{Ability}_i$. We normalize Ability_i to equal zero at the 90$^{\text{th}}$ percentile, and so that one unit corresponds to a one percentile point increase in ability.

The regression includes each applicant's vector of subject scores on the ICFES exam, which we denote by X_i. We interact these subject scores with dummies for programs m and application cohorts t. Since flagship admission is solely determined by ICFES subject scores, the coefficients β and β^A are identified from changes in the probability of admission across programs and cohorts conditional on X_i.

Columns (A)–(D) present reduced form estimates of β and β^A from the above equations. Columns (E)–(F) present analogous estimates from 2SLS regressions, where we replace Admitted_{imt} with an indicator for enrolling in the flagship program and use Admitted_{imt} as an instrument for flagship enrollment.

The sample in columns (A)–(C) and (E) includes all applicants to flagship STEM programs. In columns (D) and (F), the sample includes the subset of applicants with formal sector earnings in 2017.

Parentheses contain standard errors clustered at the program/cohort level.

Table A10. Number of students admitted to non-STEM programs

Quota variation	Program	Aug 1999	Jan 2000	Aug 2000	Jan 2001	Aug 2001	Jan 2002	Aug 2002	Jan 2003	Aug 2003	Jan 2004
		Number admitted by semester of application									
Tracking admissions	Accounting (day)		25	**97**		**194**		**178**		**96**	
	Accounting (night)			**99**		**101**		**95**		**93**	
	Architecture	49	35	**102**		**125**		**100**		**132**	
	Business (day)	51		**106**		**196**		**184**		**100**	
	Business (night)		48	**105**		**103**		**89**		**90**	
	Foreign Trade							54		**92**	
Minimal	Other programs (mean)	38	12	51	39	50	42	48	40	47	60

Notes: This table shows the number of students in our sample who were admitted to the flagship's non-STEM programs. Columns denote the semester of application, which we observe from August 1999 to January 2004.

The first six rows show programs that used "tracking" admissions for the cohorts indicated in **bold** (see Section 5.2). The last row shows the mean number of admits for the other 24 non-STEM programs in our sample.

Bold numbers are cohorts that we define as having large quotas for our binary measure of L_{mt} (see Section 5.3).

Table A11. Effects of quota variation on returns to non-STEM flagship programs

	(A)	(B)	(C)	(D)	(E)
	RD estimates			RD diff-in-diff coefs	
Dependent variable	Large quota cohorts	Small quota cohorts	p value diff	Effect of large quota	Effect of 60 extra admits
Ability percentile below threshold	0.701	0.881		−0.114	−0.044
				(0.025)	(0.020)
Panel A. Reduced form regressions					
Enrolled in flagship program	0.715	0.582	0.590	0.080	−0.029
	(0.048)	(0.260)		(0.185)	(0.068)
Completed flagship program	0.322	0.260	0.803	0.133	−0.083
	(0.053)	(0.262)		(0.228)	(0.088)
Employed in formal sector in 2017	0.090	−0.108	0.039	−0.023	−0.000
	(0.051)	(0.081)		(0.133)	(0.062)
Log monthly earnings in 2017	−0.220	0.343	0.029	−0.287	−0.168
	(0.133)	(0.220)		(0.246)	(0.230)
Panel B. 2SLS regressions					
Flagship graduation rate	0.451	0.447	0.987	−0.131	−0.086
	(0.058)	(0.234)		(0.059)	(0.077)
Return to flagship enrollment	−0.324	0.503	0.002	−0.327	−0.176
	(0.179)	(0.201)		(0.386)	(0.325)
N	1,055	187		7,664	7,664

Notes: This table shows how returns to non-STEM flagship programs changed with the size of the admission quotas.

Columns (A)–(C) present RD coefficients from separate regressions for cohorts with large and small quotas. The sample for column (A) includes the cohorts in bold for the six programs listed in Appendix Table A10. The sample for column (B) includes the non-bold cohorts of these six programs. We weight observations in these regressions so that the sample in each column has the same distribution of programs as in the combined sample for both columns. Column (C) displays the p value from an F test that the RD coefficients are equal for the two cohort groups.

Columns (D)–(E) present estimates from our RD differences-in-differences specification. The sample includes applicants to all non-STEM programs. In column (D), the variable of interest, L_{mt}, is an indicator for the cohorts in bold in Appendix Table A10. In column (E), we define L_{mt} as the total number of admits in each program/cohort divided by 60.

In the first row, columns (A)–(B) show the mean graduation-anchored ability percentile for applicants 1–5 positions below the admission thresholds, and columns (D)–(E) show estimates of π from equation (5) with mean ability below the threshold as the dependent variable. Panel A displays reduced-form estimates of θ from equation (2) (columns (A)–(C)) and of π from equations (4)–(5) (columns (D)–(E)) using the dependent variable in the row header. Panel B displays 2SLS estimates of β from equation (3) (columns (A)–(C)) and of the RD differences-in-differences coefficient (columns (D)–(E)) using flagship program completion and log monthly earnings in 2017 as dependent variables. All regressions include only applicants within 30 positions of the admission thresholds.

Parentheses contain standard errors clustered at the program/cohort level.

Table A12. Effects of flagship STEM admission on mean earnings in applicants' colleges and majors

	(A)	(B)	(C)	(D)	(E)
	Lower-ability applicants		High-ability applicants		
Dependent variable	Mean below threshold	RD coef	Mean below threshold	RD coef	p value diff
Log monthly earnings in 2017	14.132	0.196 (0.069)	14.318	0.042 (0.065)	0.101
Mean log earnings at college	14.055	0.041 (0.010)	14.104	0.005 (0.009)	0.008
Mean log earnings in major	14.086	0.096 (0.012)	14.151	0.041 (0.011)	0.001
Mean log earnings in college/major	14.084	0.118 (0.013)	14.158	0.056 (0.013)	0.001
N	303	3,390	354	3,309	

Notes: This table displays RD coefficients from separate regressions for lower- and high-ability applicants, which we define as in Table 4. Columns (A)–(B) include lower-ability STEM applicants, and columns (C)–(D) include high-ability STEM applicants. All regressions include only applicants within 30 positions of the admission thresholds.

The dependent variable in the first row is log monthly earnings in 2017, which replicates the results from Table 4. In the bottom three rows, the dependent variables are the mean log earnings in the college, major, or college/major pair that an applicant enrolled in. We calculate mean earnings using all individuals in our national administrative data (see Section 1.2). We define majors using the Ministry of Education's categorization of college programs into 55 different fields of study. For individuals who did not enroll in college, we use the mean log earnings for all non-enrollees in our administrative data.

Columns (A) and (C) display means of the dependent variables listed in the row header for applicants 1–5 positions below the admission thresholds. Columns (B) and (D) display reduced-form RD coefficients, θ, from equation (2). Column (E) displays the p value from an F test that the RD coefficients are equal for the two ability groups.

Parentheses contain standard errors clustered at the individual level.

B. Theoretical and empirical appendix

B.1. **Framework details.** This appendix contains a full derivation of our framework from Section 3, which describes how the returns to enrolling in a flagship STEM program can vary with ability.

We consider a population of high school graduates indexed by i with pre-college ability α_i. Students can choose from a large number of college programs $p \in P$, where programs are defined by both an institution and a field of study. The set P also includes the option of not enrolling in college at all, which we denote by $p = 0$.

We define the following potential outcomes that describe an individual's returns to enrolling in each program:

- Let v_{ip}^e represent individual i's potential skill value added from *enrolling* in program p. This term reflects, for example, the skills an individual learns in first-year courses.
- Let g_{ip} denote individual i's potential graduation outcome in program p. In other words, $g_{ip} = 1$ for individuals who would successfully complete the program if they enrolled, and $g_{ip} = 0$ for individuals who would drop out.
- Let v_{ip}^g represent the additional skill that individual i would gain if they *graduate* from program p.

We assume $v_{ip}^e \geq 0$ and $v_{ip}^g \geq 0$ for all p, and that $v_{i0}^e = v_{i0}^g = 0$ for the option of not attending college. Importantly, each of these three potential outcomes can depend on an individual's ability, α_i.

After college, individuals enter a competitive labor market and earn a wage equal to their skill. Under the above assumptions, individual i's potential log wage from enrolling in program p is given by:

$$w_{ip} = \alpha_i + v_{ip}^e + g_{ip} v_{ip}^g. \tag{B1}$$

An individual's wage is equal to $\alpha_i + v_{ip}^e + v_{ip}^g$ if they complete program p, and it is equal to $\alpha_i + v_{ip}^e$ if they drop out of the program.

Our empirical estimates pertain to a population of "compliers" for a flagship STEM program that we denote by s. By "compliers," we mean a group of students who would enroll in program s if and only if they are offered admission. If these students are not admitted, they enroll in their next-choice program that we denote by $c(i) \in P$. Next-choice programs can vary across individuals in the complier group, and they may differ from program s in institution and/or field of study.

We begin by examining the average wage returns to *enrolling* in the flagship STEM program in Section 4. We denote this return by $E[w_{is} - w_{i,c(i)}]$, where where the expectation is defined over all compliers who are close to the flagship's admission threshold. Using the wage equation (B1) and that fact that g_{ip} is binary, this return is given by:

$$E[w_{is} - w_{i,c(i)}] = E[v^e_{is} - v^e_{i,c(i)}] + \Big\{E[v^g_{is}|g_{is} = 1] - E[v^g_{i,c(i)}|g_{i,c(i)} = 1]\Big\}E[g_{is}] + E[v^g_{i,c(i)}|g_{i,c(i)} = 1]E[g_{is} - g_{i,c(i)}] \tag{B2}$$

Our results in Section 5 show how the returns to enrolling in a flagship STEM program vary with pre-college ability. In notation this estimand is $dE[w_{is} - w_{i,c(i)}|\alpha_i = \alpha]/d\alpha$—the change in the mean wage return to program s from an increase in ability, α. Using equation (B2) and letting $E_\alpha[x] \equiv E[x|\alpha_i = \alpha]$ denote the expected value of a variable x conditional on ability level $\alpha_i = \alpha$, this term is given by:

(B3)

$$\frac{dE_\alpha[w_{is} - w_{i,c(i)}]}{d\alpha} = \underbrace{\frac{dE_\alpha[v^e_{is} - v^e_{i,c(i)}]}{d\alpha}}_{\text{Term 1}} + \underbrace{\frac{dE_\alpha[v^g_{is}|g_{is} = 1]}{d\alpha}E_\alpha[g_{is}]}_{\text{Term 2}} - \underbrace{\frac{dE_\alpha[v^g_{i,c(i)}|g_{i,c(i)} = 1]}{d\alpha}E_\alpha[g_{i,c(i)}]}_{\text{Term 3}}$$

$$+ \underbrace{\Big\{E_\alpha[v^g_{is}|g_{is} = 1] - E_\alpha[v^g_{i,c(i)}|g_{i,c(i)} = 1]\Big\}\frac{dE_\alpha[g_{is}]}{d\alpha}}_{\text{Term 4}} + \underbrace{E_\alpha[v^g_{i,c(i)}|g_{i,c(i)} = 1]\frac{dE_\alpha[g_{is} - g_{i,c(i)}]}{d\alpha}}_{\text{Term 5}}$$

Our RD analysis of the returns to the flagship STEM programs yields three main results. First, there is a positive mean earnings return to enrolling in these STEM programs for marginal admits (Table 3). Second, STEM graduation rates increase with ability (Tables 4 and 6), and the effect of enrollment on the probability of earning any college degree also weakly increases with ability (Table 7). Third, mean earnings returns to enrolling in these programs *decrease* with ability (Tables 4 and 6).

These results lead us to explore the mechanisms through which lower-ability students can have larger earnings returns to selective STEM programs. All else equal, earnings returns increase with the probability of graduating, but there are three reasons why returns can be larger for less able students despite lower graduation rates. We summarize these three mechanisms in the following proposition.

Proposition. *Suppose that:*

(i) The skill return to graduating from program s is non-negative for all ability levels,

$$E_\alpha[v^g_{is}|g_{is} = 1] - E_\alpha[v^g_{i,c(i)}|g_{i,c(i)} = 1] \geq 0;$$

(ii) Both absolute and relative graduation rates in program s are increasing in ability,

$$\frac{dE_\alpha[g_{is}]}{d\alpha} > 0 \quad \text{and} \quad \frac{dE_\alpha[g_{is} - g_{i,c(i)}]}{d\alpha} > 0.$$

Then if the wage return to enrolling in program s is <u>*decreasing*</u> *in ability,* $dE_\alpha[w_{is} - w_{i,c(i)}]/d\alpha < 0$*, at least one of the following conditions must hold:*

(a) There is a skill return to enrolling in program s that decreases with ability,

$$\frac{dE_\alpha[v^e_{is} - v^e_{i,c(i)}]}{d\alpha} < 0;$$

(b) Lower-ability students choose counterfactual programs with less degree value added,

$$\frac{dE_\alpha[v^g_{i,c(i)} | g_{i,c(i)} = 1]}{d\alpha} > 0;$$

(c) Lower-ability students have greater value added to a degree from program s,

$$\frac{dE_\alpha[v^g_{is} | g_{is} = 1]}{d\alpha} < 0.$$

This proposition follows from inspection of equation (B3). Conditions (i) and (ii) ensure that Terms 4 and 5 are non-negative since $E_\alpha[v^g_{i,c(i)} | g_{i,c(i)} = 1] \geq 0$. Mechanisms (a)–(c) determine the sign of Terms 1–3 since $E_\alpha[g_{is}] \geq 0$ and $E_\alpha[g_{i,c(i)}] \geq 0$.

We explore the empirical evidence on these three mechanisms in Section 6.

B.2. **Data and merging.** This section provides details on our data sources and merging.

Our base dataset includes lists of all applicants to Universidad del Valle's undergraduate programs from Fall 1999 to Spring 2004. These data were provided by Univalle, and they include the program/cohort that applicants applied to, their admission scores, and their admission decisions.

We combine the flagship application records with three individual-level administrative datasets provided by the Colombian government. The first dataset includes records from Colombia's national standardized college entrance exam, which was formerly called the ICFES exam and is now called *Saber 11*. The data were provided by the agency that administers the exam, and it contains all students who took the exam between 1998–2004. The ICFES exam is also used by the Colombian government for high school accountability, so it is taken by nearly every high school graduate in the country. The main variables of interest are individuals' scores on each exam subject and demographic characteristics.

The second administrative dataset includes enrollment and graduation records from the Ministry of Education. These records include the institution, program of study, and graduation outcome for students who enrolled in college between 1998–2012. The Ministry's records cover almost all colleges in Colombia, although it omits a few schools due to their small size or inconsistent reporting. To describe the set of colleges that are included in the

Table B1. Higher education institutions in the Ministry of Education records

	(A) Number of colleges	(B) Number of exit exam takers/year	(C) Prop. of colleges in records
University	122	134,496	1.00
University Institute	103	53,338	0.88
Technology School	3	2,041	1.00
Technology Institute	47	15,092	0.82
Technical/Professional Institute	35	11,408	0.99
Total	310	216,375	0.96

Notes: Column (A) depicts the number of colleges that have *Saber Pro* exit exam takers in 2009–2011 using administrative records from the testing agency. Colleges are categorized into the Ministry of Education's five higher education institution types. Column (B) shows the number of 2009–2011 exam takers per year. Column (C) shows the proportion of colleges that appear in the Ministry of Education records, where colleges are weighted by the number of exit exam takers.

Ministry of Education records, we use another administrative dataset from a college exit exam called *Saber Pro* (formerly ECAES). This national exit exam is administered by the same agency that runs the ICFES college admission exam, and it became a requirement for graduation from any higher education institution in 2009. Column (A) in Table B1 depicts the 310 colleges that have any exit exam takers in these administrative records in 2009–2011. These colleges are categorized into the Ministry of Education's five types of higher education institutions, which are listed in descending order of their on-time program duration.[28] Column (B) shows the number of exit exam takers per year. The majority of exam takers are from university-level institutions, with fewer students from technical colleges. Column (C) shows the fraction of these 310 colleges that appear in the Ministry of Education records that we use in our analysis. These proportions are weighted by the number of exam takers depicted in column (B). Column (C) shows that the Ministry of Education records include all universities but are missing a few technical colleges.[29] Overall, 96 percent of exit exam takers attend colleges that appear in the Ministry of Education records.

The third administrative dataset includes earnings records collected by the Ministry of Social Protection. The records are from the Ministry's electronic tax record system called *Planilla Integrada de Liquidación de Aportes* (PILA). Our data include monthly earnings in 2017 for any individual who worked at a firm that was registered with the Ministry. Our main income measure is average monthly earnings, which we compute by dividing total

[28] Most programs at universities require 4–5 years of study, while programs at Technical/Professional Institutes typically take 2–3 years.

[29] The largest omitted institutions are the national police academy (*Dirección Nacional de Escuelas*) and the Ministry of Labor's national training service (*Servicio Nacional de Aprendizaje*).

annual earnings by the number of employment months in 2017. We also use an indicator for appearing in the PILA dataset as a measure of formal employment.

We merge the flagship application data into the ICFES data using applicants' full names. Since the ICFES exam is required for admission to Univalle, most applicants appear in the ICFES administrative dataset. Most individuals match uniquely on name, but in cases with duplicate names we use information on ICFES exam cohort and high school location to identify the correct match.[30] Through this process, we are able to match 84 percent of individuals in the flagship application data to the ICFES records, as shown in columns (A)–(B) in Table B3 below. The vast majority of non-matches occur because individuals took the ICFES exam prior to 1998, when our records begin.[31]

We merge the ICFES and Ministry of Education datasets using individuals' national ID numbers, birth dates, and names. We define a match from this merge as observations that have either: 1) the same ID number and a fuzzy name match; 2) the same birth date and a fuzzy name match; or 3) an exact name match for a name that is unique in both records.[32] 39 percent of the 1998–2004 ICFES exam takers appear in the Ministry of Education records, which is comparable to the higher education enrollment rate in Colombia during the same time period.[33] A better indicator of merge success is the percentage of college enrollees that appear in the admission exam records because all domestic college students must take the exam. We match 91 percent of enrollees who took the admission exam between 1998 and 2004.[34]

Lastly, the combined dataset from the above merges was matched to the PILA earnings records by the Colombian statistical agency *Departamento Administrativo Nacional de Estadística* (DANE). DANE also merged these datasets using national ID numbers, names, and birth dates. The fraction of individuals in the 1998–2004 ICFES exam cohorts who were matched to the 2017 earnings dataset is 56 percent. To benchmark this merge rate, we

[30] If there are duplicates, we select the individual who took the ICFES exam prior to Univalle application, and who attended a high school in the Valle del Cauca region. If these criteria do not identify a unique ICFES exam taker, we consider the applicant to be a non-match.

[31] Many Colombians wait a year or more after high school before applying to college.

[32] Nearly all students in these records have national ID numbers, but Colombians change ID numbers around age 17. Most students in the admission exam records have below-17 ID numbers (*tarjeta*), while most students in the college enrollment and earnings records have above-17 ID numbers (*cédula*). Merging using ID numbers alone would therefore lose a large majority of students.

[33] The gross tertiary enrollment rate ranged from 23 percent to 28 percent between 1998 and 2004 (World Bank World Development Indicators, available at: http://data.worldbank.org/country/colombia). This rate is not directly comparable to our merge rate because not all high school aged Colombians take the ICFES exam. About 70 percent of the secondary school aged population was enrolled in high school in this period. Dividing the tertiary enrollment ratio by the secondary enrollment ratio gives a number roughly comparable to our 39 percent merge rate.

[34] Approximately 16 percent of students in the Ministry of Education records have missing birth dates, which accounts for most of the non-matches.

use Colombian household survey data (GEIH) on individuals in the 1981–1987 birth cohorts with at least a high school degree. In this population, the fraction of individuals who worked and had a contract for their employment was also 56 percent in 2017. This suggests that the DANE merge identified nearly all individuals in our sample with formal sector jobs.

B.3. **Analysis sample.** This section provides details on the sample we use for our analysis.

Our sample includes all of Univalle's bachelor's degree programs where we can identify the effects of admission. Our initial dataset includes applicants to 74 different degree programs from Fall 1999 to Spring 2004. We exclude 26 of these programs from our sample for one of two reasons, as shown in Table B2. First, we exclude technical/professional programs to focus on bachelor's degree attainment (column (C)). Second, we exclude programs with fewer than two cohorts in which any applicant was rejected (column (E)), which is necessary for our RD differences-in-differences design. Excluded programs tend to attract fewer applicants and were offered only a few times during our data period. Our sample includes the remaining 48 degree programs listed in Appendix Table A1.

Table B3 shows the applicants to these 48 programs that we include in our sample. Column (A) shows that our initial dataset includes 20,001 applicants to the STEM programs in our sample (Panel A) and 29,041 applicants to other programs (Panel B). We exclude applicants for the three reasons shown in columns (B)–(D) of Table B3. First, we drop applicants who do not appear in our ICFES dataset (column (B)), as described in Section B.2. Second, we exclude applicants in special disadvantaged admission groups who were not subject to the flagship's primary admission thresholds (column (C)). During this time period, Univalle maintained special admission quotas for disabled, indigenous, and military applicants. Third, we drop applicants from cohorts where no applicants were rejected (column (D)), which is necessary for our RD strategy. After these restrictions, our sample includes 16,022 STEM applicants and 23,439 applicants to other programs.

Most of our regressions focus on the subset of applicants whose admission scores are within h ranks of the tracking threshold. Our benchmark model uses $h = 30$, which is roughly the mean of the Calonico et al. (2014) bandwidths across all dependent variables. Column (F) shows that this RD sample includes 6,699 STEM applicants and 7,664 applicants to other programs. Applicants in our RD sample tend to have higher pre-college ability than those in the full sample. In addition, these applicants come from slightly more advantaged socioeconomic backgrounds, and are less likely to identify as female.

TABLE B2. Programs excluded from sample

(A) Faculty area	#	(B) Program	(C) Degree level	(D) Application cohorts	(E) Cohorts with rejects	(F) Total applied
Engineering	1	Environmental Management	Technical	4	2	538
	2	Food Science	Technical	1	1	195
	3	Forest Protection	Technical	1	0	12
	4	Information Systems	Technical	1	1	201
	5	Soil and Water Conservation	Technical	4	2	253
Health	6	Prehospital Care	Technical	4	2	3,014
Humanities	7	Geography	Bachelor's	2	1	97
	8	Philosophy	Professional	2	2	106
	9	Physical Education	Professional	2	1	316
	10	Political Studies	Bachelor's	3	1	337
	11	Recreation (night)	Bachelor's	2	1	112
	12	Teaching (Biology & Chemistry)	Bachelor's	2	0	44
	13	Teaching (Elem. Math, day)	Bachelor's	1	0	34
	14	Teaching (Elem. Math, mixed)	Bachelor's	1	0	30
	15	Teaching (Elem. N. Science, day)	Bachelor's	1	1	138
	16	Teaching (Elem. N. Science, mixed)	Bachelor's	1	0	13
	17	Teaching (Math & Physics, day)	Bachelor's	1	1	65
	18	Teaching (Math & Physics, mixed)	Bachelor's	1	0	18
	19	Teaching (Modern Languages, day)	Bachelor's	1	1	39
	20	Teaching (Modern Languages, night)	Bachelor's	1	0	37
	21	Teaching (Phys. Ed. & Health)	Bachelor's	2	1	111
	22	Teaching (Physical Education, day)	Bachelor's	1	1	55
	23	Teaching (Physical Education, mixed)	Bachelor's	2	0	43
	24	Teaching (Physical Math)	Bachelor's	1	0	23
	25	Teaching (Popular Education)	Bachelor's	1	1	45
Integrated arts	26	Music	Bachelor's	1	1	110
Total			Bachelor's	44	21	5,986

Notes: Columns (A)–(B) list the Univalle programs that we exclude from our sample and their faculty areas at the university. Column (C) reports the program's degree level (technical, professional, or bachelor's). Column (D) shows the total number of application cohorts from August 1999 to January 2004. Column (E) shows the number of cohorts during this period in which any applicant was rejected. Column (F) shows the total number of applicants during this period.

Table B3. Analysis sample

	(A)	(B)	(C)	(D)	(E)	(F)
		Excluded applicants				
	All applicants	Missing ICFES scores	Special admission group	No rejected applicants	Full sample	RD sample
Panel A. STEM applicants						
Ability percentile	0.783		0.811	0.838	0.780	0.842
Age	18.713		19.544	18.997	18.686	18.947
College educated father	0.426		0.355	0.454	0.427	0.440
College educated mother	0.361		0.330	0.358	0.361	0.373
Family income > 2x min wage	0.576		0.470	0.612	0.576	0.599
Female	0.357		0.274	0.312	0.360	0.319
N	20,001	3,077	310	592	16,022	6,699
Panel B. Other applicants						
Ability percentile	0.735		0.778	0.846	0.733	0.810
Age	18.923		19.735	20.596	18.879	19.353
College educated father	0.408		0.431	0.370	0.408	0.424
College educated mother	0.344		0.375	0.300	0.344	0.354
Family income > 2x min wage	0.560		0.498	0.609	0.560	0.589
Female	0.637		0.539	0.541	0.641	0.588
N	29,041	4,746	462	394	23,439	7,664

Notes: Column (A) shows the total number of applicants to the 48 flagship programs in our sample (see Appendix Table A1). Column (B) shows the number of applicants who do not appear in the ICFES dataset. Column (C) lists the number of students who were admitted through special quotas for disadvantaged groups. Column (D) shows the number of applicants to program/cohort pairs in which no applicants were rejected. Column (E) shows our full analysis sample, which is equal to column (A) minus the applicants in columns (B)–(D). Column (F) shows the subset of applicants from column (E) who are within 30 positions of the admission threshold in their application pool.

Panel A includes applicants to the flagship's STEM programs, and Panel B includes applicants to other programs. Demographic characteristics are not reported in column (B) because they are based on variables in the ICFES dataset.