Heuristic Self-Evaluation and the Returns to STEM

Yoav Goldstein*
24 September 2025

Abstract

Human capital decisions shape long-term outcomes, yet students often make them under uncertainty about their own ability. Admissions to high-earning STEM programs rely on test scores that are noisy signals of ability. This paper studies how heuristics and self-evaluation influence the decision to pursue such programs. Using linked administrative data, I implement a regression discontinuity at a round-number threshold on an initial university entrance test. Although the threshold has no official role in admissions, crossing it sharply increases applications to STEM, consistent with left-digit bias: students interpret a higher first digit as a signal of ability. Affected students retake the exam, improve their scores, and enroll in STEM. Over the next two decades, they are more likely to work in the technology sector, earn higher wages, and reach the upper tail of the earnings distribution. The findings show how cognitive heuristics and uncertainty can deter students from selective fields, generating persistent disparities in economic opportunity unrelated to actual talent.

JEL Codes: I26, J24, J31, D91

*The Eitan Berglas School of Economics, Tel Aviv University, yoavg2@mail.tau.ac.il. I thank the Israeli Central Bureau of Statistics for providing access to the data I use in this study at its protected research room in Tel Aviv. I am grateful to my supervisor, Analia Schlosser, for her invaluable guidance and support throughout the research process. I also thank Hadar Avivi, Netanel Ben Porath, Zach Bleemer, Oren Danieli, Naomi Gershoni, Josh Goodman, Matan Kolerman, Victor Lavy, Dor Leventer, Ro'ee Levy, Itay Saporta-Eksten, Ity Shurtz, Dan Zeltzer, as well as participants at various conferences and seminars for useful comments and suggestions. I thank the Azrieli Foundation and the Pinhas Sapir Economic Policy Forum for financial support. An earlier version of this paper has been circulated under the name "Learning from Round Test Scores: University Field Choices and Career Outcomes".

1 Introduction

Human capital investments are among the most consequential economic decisions, yet they are often made under uncertainty about one's own ability (Altonji, 1993). In modern economies, students sort into fields with sharply different wage returns (Kirkeboen et al., 2016). Admission to high-return fields such as STEM typically relies on a test-based meritocracy, designed to align opportunities with ability. I examine how uncertainty and cognitive heuristics distort degree choice and track their long-term effects.

I use administrative data and a regression discontinuity to study this question. In test-based admissions, scores are noisy signals of fit. To generate quasi-experimental variation in perceptions, I exploit left-digit bias—the tendency to place greater weight on the leftmost digit of a number (e.g., Strulov-Shlain, 2023). I compare students whose first university entrance test score falls just above versus just below a round number high in the score distribution but below admission cutoffs for selective STEM fields. This quasi-random boost in perceived success does not change immediate admissions, but crossing the threshold leads some students to retake the test, improve their scores, and enter high-earning STEM programs.

This raises the question of whether such heuristic boosts merely reflect overconfidence or actually change students' long-term paths. Evidence points to real gains. Years later, these students are more likely to work in technology firms, earn much higher wages, and appear among the top earners in their cohorts. Taken together, the results show that uncertainty about one's own ability—combined with heuristic interpretations of noisy scores—can deter entry into high-return fields and cause capable students to undermatch.

This research relies on an administrative dataset from the Israeli Central Bureau of Statistics. As detailed in Section 2, the dataset links students' entrance test scores with records on university applications, degree enrollment and completion, and subsequent labor-market outcomes. The analysis focuses on individuals who first took the test between 2000 and 2009 at age 20 or younger.[1] Tax records extend through 2023, allowing me to track careers for up to two decades after the initial test.

In Israel, tech employment is associated with much higher wages at all education levels. Among university graduates, those with degrees in computer science or electrical engineering (hereafter "selective STEM") earn about 50–200 percent more than graduates of other fields. By their early thirties, the median selective STEM graduate earns more than 90 percent of all workers, across the entire national earnings distribution. Admission to Israeli universities is based on a composite of the entrance test score and the matriculation GPA, typically

[1] Older test-takers, common due to mandatory military service, typically take the test just before applying to university, leaving little scope for behavioral responses. Results for this group are reported in the appendix.

weighted equally and considering only the highest scores. Decisions rely exclusively on these quantitative measures, with no subjective evaluations. Admissions are decentralized: each program (field by institution) sets its own cutoff, and prior-year thresholds are published online. Requirements vary across programs, with selective STEM among the most demanding. Application, testing, and tuition costs are relatively low.

Section 3 outlines the empirical approach, which exploits a regression discontinuity design around a round-number threshold on the first entrance test. The threshold is relatively high—only about 30 percent of test-takers score above it—yet it is below the admission cutoff for selective STEM degrees. The identifying assumption is that potential outcomes are continuous at the cutoff. This is supported by both institutional features and the data: score distributions and baseline characteristics are smooth at the threshold.

Estimates presented in Section 4 show that crossing the round-number threshold increases applications to selective STEM degrees by 3.8 percentage points within five years of the test, from a baseline of 9.6 percent.[2] This effect is persistent and reflects a shift in field choice rather than an increase in total applications or greater institutional selectivity.

Because the administrative data lack direct measures of self-evaluation, I examine alternative explanations. Admission probabilities do not change discontinuously at this threshold. Nor is the pattern consistent with students mistakenly treating 600 as a formal cutoff: most STEM applicants retake the test and raise their scores substantially before applying, indicating they know 600 is insufficient for admission. These patterns point to heuristic self-evaluation as the main driver. At the same time, round numbers may act as salient signals within students' social circles. In line with this, younger siblings are more likely to take the test if an older sibling scores just above the threshold.

Heterogeneity analysis shows that both men and women increase applications to selective STEM degrees when crossing the threshold, with a larger effect for men. The strongest heterogeneity, however, arises from quantitative preparation. The test includes three domains—quantitative, verbal, and English—and students in selective STEM programs typically excel in the quantitative domain. Splitting by this dimension, I find that students with strong quantitative scores are the most likely to increase applications to selective STEM fields. By contrast, students without a quantitative advantage respond differently: instead of applying more, they are less likely to retake the test. This reflects another behavioral response to

[2]The appendix shows results for the highest round-number threshold as well. At lower thresholds, STEM applications are too rare to detect effects. At higher thresholds, application rates are already high, and the bias instead appears in behaviors such as pursuing advanced degrees, including in selective STEM fields.

round numbers—treating them as goals—so that surpassing a salient threshold increases satisfaction with achievement (Pope and Simonsohn, 2011).[3]

Examining degree enrollment and completion, I find that about half of the marginal applicants ultimately enroll in selective STEM degrees. Among those who do not gain admission, many shift toward other high-earning STEM programs. Overall, the probability of ever enrolling in one of Israel's most rewarding degrees—all of which are in STEM—rises by 3.8 percentage points from a baseline of 22.0 percent.

Tracking these students into the labor market reveals substantial benefits. By 2022–2023, when cohorts are aged 30–44, those scoring just above the threshold are 3.5 percentage points more likely to work in technology firms. Their annual earnings rise by 13,200 NIS (3,500 USD or 5 log points), and monthly wages by 1,400 NIS (400 USD or 4 log points). These gains emerge roughly ten years after the test and persist for at least another decade, consistent with sustained returns to human capital investment. Above-threshold students are also more likely to reach the top of the earnings distribution. For example, the probability of being in the top 5 percent increases by 1.6 percentage points from a baseline of 6.8 percent.

Section 5 quantifies the causal returns to STEM choice. Unlike conventional regression discontinuity designs that exploit formal admissions thresholds (e.g., Bleemer and Mehta, 2022), this analysis leverages quasi-random variation in application behavior generated by heuristic-based self-evaluation. Identification requires a strong exclusion restriction: scoring just above the round number must affect outcomes only through the decision to pursue selective STEM degrees. I provide evidence supporting this assumption and focus on the subsample with strong quantitative preparation, where violations are less likely. Results imply large returns. By ages 30–32, monthly wages are higher by about 20,000 NIS, corresponding to an increase of roughly 120 percent in the log-wage specification. I also report more conservative estimates that relax the exclusion restriction, which still yield significant gains of about 12,500 NIS per month.

This paper relates to the literature on human capital decisions. A longstanding literature examines how beliefs about one's own ability shape these choices (Altonji, 1993; Zafar, 2011; Arcidiacono et al., 2012; Stinebrickner and Stinebrickner, 2012; Wiswall and Zafar, 2015; Papay et al., 2016; Smith et al., 2017; Avery et al., 2018; Li and Xia, 2024; Arcidiacono et al., 2025).[4] The results here contribute to the literature in four ways. First, they show that heuristic interpretations of test scores create behavioral frictions in the process of learning

[3]The appendix presents a simple conceptual framework that reconciles these patterns. Round-number scores generate two channels: (i) heuristic updating about ability, which encourages uncertain students to aim higher; and (ii) the satisfaction from passing a salient benchmark, which reduces the incentive to retest.

[4]Other related work shows that shocks to test scores (e.g., Diamond and Persson, 2016; Ebenstein et al., 2016; Lavy and Megalokonomou, 2019; Landaud et al., 2024) and within-class rank (Elsner and Isphording, 2017;

about one's own ability. Second, they demonstrate that uncertainty can generate undermatch in field-of-study choice. While earlier work on STEM has emphasized overmatch—students entering STEM and later switching out after negative grade shocks (Stinebrickner and Stinebrickner, 2014)—I provide evidence in the opposite direction: capable students are deterred from pursuing selective STEM degrees even though they could succeed. Third, by following students for two decades, I document substantial economic losses driven by this underestimation of ability. Fourth, I show that heuristic-based ability learning creates spillovers within families, influencing younger siblings' choices.

Additionally, this paper contributes to the literature on returns to postsecondary education. Prior work shows that field specialization leads to highly heterogeneous returns (Hastings et al., 2013; Kirkeboen et al., 2016; Altonji et al., 2016; Heinesen et al., 2024; Bleemer and Mehta, 2022; Daly et al., 2022; Andrews et al., 2024; Ng and Riehl, 2024). Focusing on STEM, my results align with evidence from Colombia (Ng and Riehl, 2024), where less-prepared students benefit more from selective STEM. In my setting, lower-scoring students who later enter selective STEM earn returns that are not smaller than those of the average student. These findings add to recent evidence from selective institutions in the United States (Bleemer, 2021, 2022; Black et al., 2023) in challenging the influential mismatch hypothesis, which posits that lower-scoring students are harmed in selective programs.

An important distinction from prior work is that I estimate returns at the application margin, exploiting quasi-random perception boosts that shift application behavior. This design targets uncertain students. The policy implications also differ: while admission-margin estimates inform the benefits of expanding program slots, my results highlight another lever—encouraging capable but hesitant students to apply. However, my results suggest that this lever has a limited scope. Because self-evaluation effects concentrate among students already considering STEM, encouragement is unlikely to substantially reduce gender or socioeconomic gaps. Closing these gaps likely requires earlier pipeline policies—such as strengthening quantitative preparation—rather than belief adjustments alone.

Finally, the paper contributes to the heuristics and biases literature. Prior work documents left-digit bias in many settings, including consumer demand and firm pricing (Lacetera et al., 2012; Strulov-Shlain, 2023), firm wage-setting (Dube et al., 2025), expert judgments (Olenski et al., 2020; Shurtz, 2022), and goal setting (Pope and Simonsohn, 2011). Goal-setting behavior is also documented in SAT retaking in the United States (Pope and Simonsohn, 2011; Goodman et al., 2020). I document similar effects in Israel but emphasize an additional channel: self-evaluation. Evidence from China also points to a similar mechanism, as

Murphy and Weinhardt, 2020; Elsner et al., 2021; Denning et al., 2023) also affect education and careers, possibly through the same mechanisms emphasized here, but not exclusively.

students become more ambitious in Gaokao applications after crossing round numbers (Li and Qiu, 2023). My main contribution is to link heuristic-based decisions in education to long-term labor-market outcomes. The results show that heuristics can shape major life choices and generate persistent economic losses. More broadly, while recent research shows that complexity or opacity in admissions systems can distort enrollment (Dynarski et al., 2021; Kapor, 2024), my findings reveal a different limitation. Even in transparent systems, uncertainty about ability combined with heuristic responses to test scores distorts decisions and creates long-term inequality unrelated to talent.

2 Data and Setting

2.1 Data Sources

This study uses an administrative database from Israel's Central Bureau of Statistics, accessed in its secure research lab. The population consists of all students enrolled in Israeli high schools (tenth grade) between 1995 and 2016. The dataset links records from several administrative sources (all detailed in Appendix B.1), most importantly: university entrance test scores, postsecondary education records, and tax data.

The University Psychometric Entrance Test (UPET). The UPET is a standardized exam that comprises quantitative, verbal, and English components; each component is scored 50–150, and the composite total is a standardized index scaled to 200–800. The exam is administered several times a year throughout Israel, it can be retaken multiple times, and universities consider only the highest total score. The database records all UPET attempts for individuals in the sample, provided by the National Institute of Testing and Evaluation. Figure A.1 presents the score distribution.[5] It can be taken in several languages (Hebrew, English, Arabic), and is relatively inexpensive, with a fixed registration fee of 560 NIS (about 150 USD). However, a large share of test-takers enroll in commercial preparation courses, which often cost about 1,000 USD.[6]

[5]Because the distribution is close to normal, very few individuals fall at the extremes. A cell-suppression policy excludes some of these cases. For analysis, I restrict the sample to total scores between 370 and 730.

[6]In 2016, 43 percent of examinees reported paying 2,000–5,000 NIS, while about one-third paid less than 2,000 NIS and roughly 30 percent paid more than 5,000 NIS; current market prices range from 3,000 to 9,000 NIS depending on course type and provider. See Knesset Research and Information Center, *The Psychometric Exam: Information on Exam Participation, Preparation Courses, and Estimated Costs*, 22 June 2022. Available at: https://fs.knesset.gov.il/globaldocs/MMM/57aeced1-28b4-ec11-814b-005056aa4246/2_57aeced1-28b4-ec11-814b-005056aa4246_11_19627.pdf (retrieved September 2, 2025).

Admission to postsecondary programs is a mechanical function of UPET scores and the GPA from Israel's Bagrut (high school matriculation) exams.[7] During our sample period, admissions did not use essays, recommendation letters, or extracurricular activities.

The analysis focuses on first UPET attempts taken between 2000 and 2009.[8] Restricting the sample to pre-2010 tests ensures long-term follow-up, with labor-market outcomes observed at least 14 years later (through 2023). Tests prior to 2000 are excluded because only a small number of students in the sample took the exam before that year and to maintain a consistent study period. This restriction reduces the sample size by about 10 percent. For each student, I retain only the first test attempt for the main analysis, while using subsequent attempts to capture retesting behavior and the maximum score achieved.

Postsecondary education records. Israel's postsecondary education system is organized as a dual structure of public universities and academic colleges. The universities are the most established, selective and research-oriented institutions. They dominate in STEM, medicine, and graduate training, and they confer almost all PhDs. By contrast, public and private colleges, which expanded rapidly in the 1990s to broaden access—particularly in peripheral regions—are generally regarded as lower-tier institutions, reflected in their admission standards, teaching orientation, and graduates' labor-market outcomes (e.g., Achdut et al., 2019).[9]

There is also meaningful heterogeneity within the university sector. According to the CBS, three institutions—Hebrew University of Jerusalem, Tel Aviv University, and the Technion—are considered the elite universities. These are typically more selective overall, though selectivity also varies substantially across fields of study within each institution.

[7]Unlike in the United States, where high school GPAs are not standardized across schools, Israel's Bagrut exams are centrally administered and graded, yielding a nationally comparable GPA. Bagrut tests are taken in grades 10–12, with students able to choose different proficiency levels, awarding one to five credit units per subject. Students may also retake these tests after high school to improve their university admission prospects. The Bagrut GPA is a weighted average over subjects in which higher proficiency levels carry larger weights. In our dataset, we observe only the final matriculation score (and proficiency level) for each individual in each subject, and scores are available in grouped ranges rather than exact values. For more details on Israel's high school and matriculation system, see, e.g., Lavy and Goldstein (2022).

[8]For tests administered between October and December, I assign them to the following calendar year, since they could not be used for admission to the same academic year, which in Israel typically begins in October.

[9]During the sample period, Israel had seven universities and dozens of colleges offering undergraduate degrees. Since then, two colleges have been recognized as universities, bringing the total to nine undergraduate universities. A tenth university offers only graduate programs. Colleges themselves are diverse: some are public and state-funded, others are private and tuition-dependent, and most focus primarily on teaching rather than research. A distinct subset consists of specialized teacher-training colleges that grant education degrees.

The system is overseen by the Israeli Council for Higher Education (CHE), the statutory body responsible for accreditation, funding, and quality assurance. All universities (and some colleges) are heavily state-funded and tuition levels are low by international standards (roughly 3,000 USD per year).

The CHE provides data on applications, enrollment, and attainment for students in our sample. Enrollment and attainment records are available for both universities and colleges, while application data during our study period are reported only for universities. This is not a major limitation, since our analysis focuses on high-earning STEM degrees, which are almost exclusively offered by universities.

Admissions are decentralized: each institution sets its own criteria and manages its own applications. Students may apply to multiple institutions and can rank up to three fields of study per institution. Degree programs (a field at a given institution) set annual admission thresholds, typically defined as a weighted average of the applicant's highest UPET score and Bagrut GPA, often with equal weights. Thresholds vary across programs and years, and are usually published online after each admission cycle.

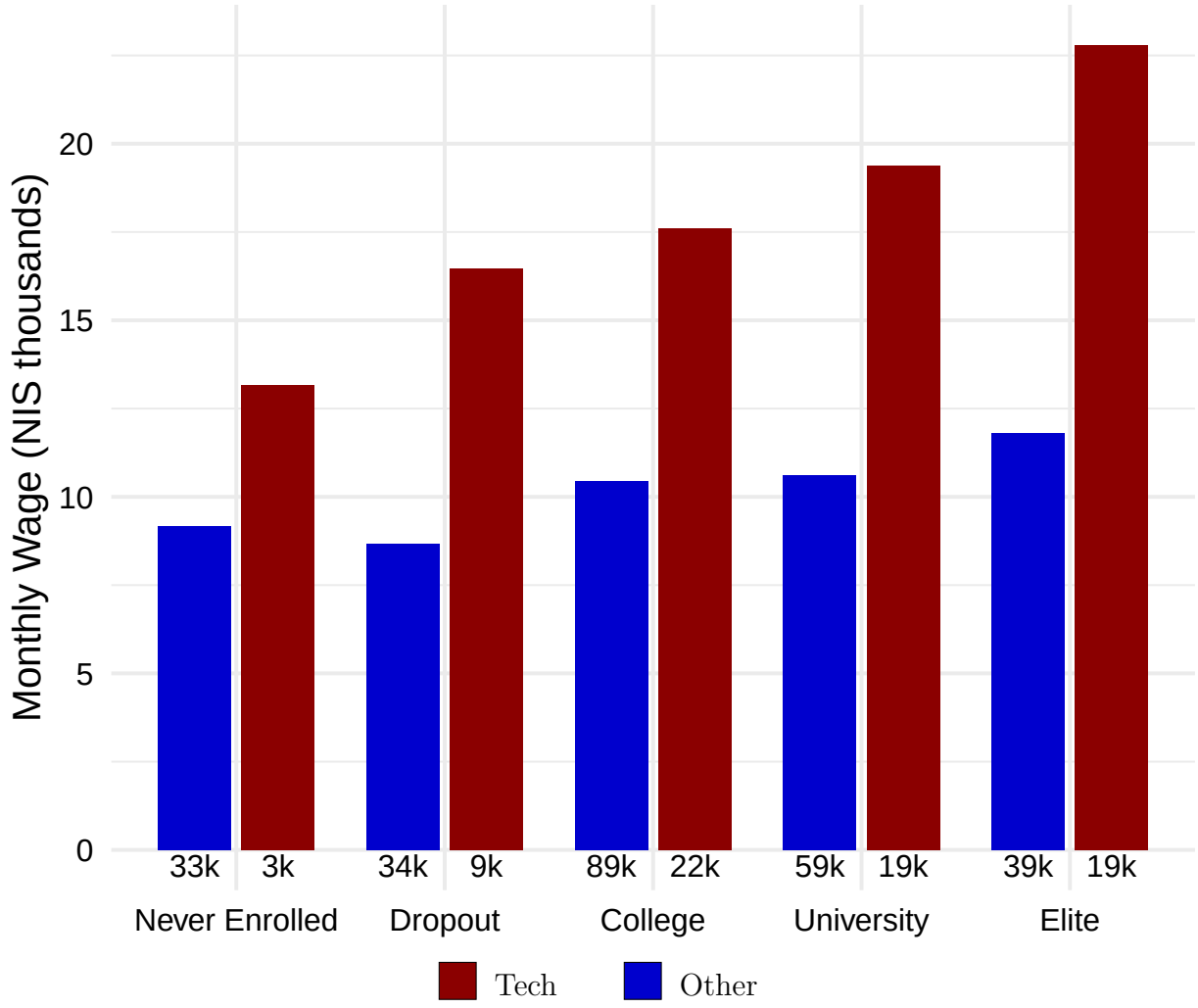
Labor-market outcomes. The dataset includes tax records for all years from 1990 to 2023, except 2021. I focus on outcomes in 2022–2023, the most recent years available, and also present age-specific trajectories and dynamic effects relative to the year of the test. From these records, I construct outcome variables including indicators for salaried employment (positive wage income and positive months worked), self-employment (positive business income), and employment in the tech sector. "Tech" follows the CBS definition covering ICT services, electronics and communications manufacturing, medical and scientific instruments, pharmaceuticals, aircraft manufacturing, and R&D services (Appendix B.1).

Additional measures include monthly wages for salaried workers, annual salaried earnings, and total annual earnings (salaried plus self-employment). I also compute age-specific earnings ranks and indicators for belonging to the top 10 percent, 5 percent, or 1 percent of the earnings distribution. These ranks are defined relative to the universe of university entrance test-takers in the same cohort. All income measures are pre-tax, and they are top-coded at the 99th percentile and restricted to observations with income above half the minimum wage. All monetary values are expressed in 2023 NIS using the CPI; USD figures use the 2023 average exchange rate (3.7 NIS to 1 USD).

The dominance of the local tech industry is illustrated in Figure 1. Across education levels, workers employed in the tech sector earn substantially higher wages in Israel. Among graduates of elite universities, the wage gap between those employed in tech and those

outside tech is almost double (23,000 versus 12,000 NIS). This gap is considerably larger than the difference between elite graduates outside tech (12,000 NIS) and individuals who never enrolled in any degree (9,000 NIS). Even individuals who never enrolled in a degree but manage to enter the tech sector earn more than elite graduates who work outside tech.

Figure 1: Average Wage by Postsecondary Education and Industry



Notes: The figure shows labor-market outcomes for test-takers employed at least one year between ages 30–32 (332,976 observations). Bars report average monthly wages (thousands of NIS, measured at ages 30–32) by postsecondary education and employment industry group. The total number of observations in each group (education × industry) is shown below each bar. Tech covers technological sectors as defined by the CBS (see Appendix B.1); "Other" includes all remaining industries. Elite refers to the three elite universities in the CBS classification (see Appendix B.1); "Universities" include all other universities. "Dropout" refers to students who enrolled in any type of institution but did not graduate.

2.2 STEM Wage Premium

As in many other countries, fields of study in Israel differ sharply in their wage premia (Krill et al., 2019; Achdut et al., 2019). Israel's thriving tech sector creates strong demand for STEM graduates. Figure 2 illustrates these patterns, showing average wages and tech-employment rates at ages 30–32 for undergraduate degree graduates. A clear pattern emerges: wages are strongly correlated with tech-employment rates. Among the common programs, computer science and electrical engineering stand out with exceptionally high wages and tech-employment shares. Graduates of these fields in all types of institutions earn more than 20,000 NIS (≈5,000 USD) per month—well above the roughly 10,000 NIS typical for most degree graduates, forming the dense cluster in the figure. They also stand out in terms of employment: about 60 percent of computer science and electrical engineering graduates work in the tech sector, compared with much lower shares in other fields. In general, all high-wage fields are also those with high rates of tech employment.

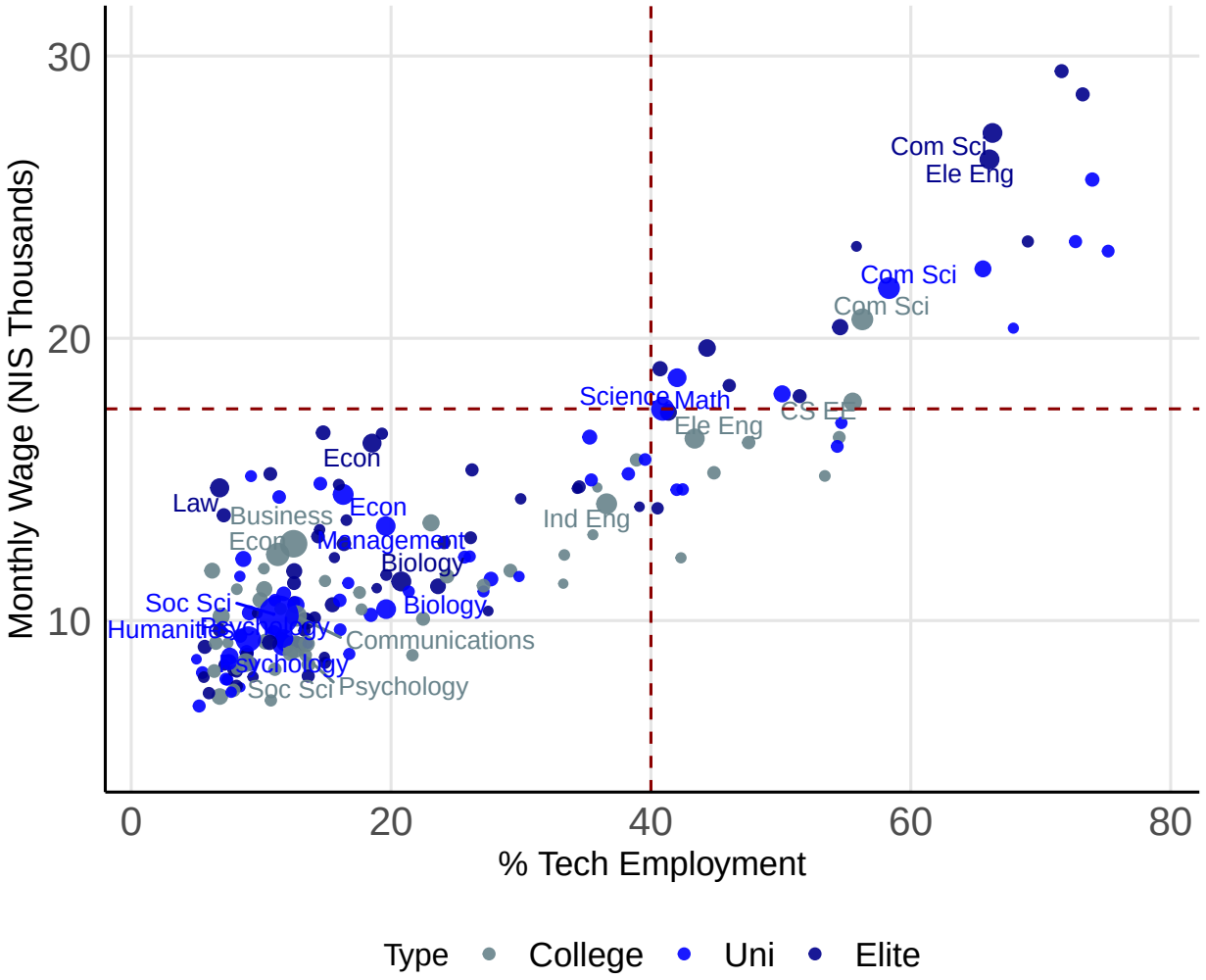
Table A.1 complements these findings by providing a breakdown of average earnings across degree types and industries. Several patterns stand out. First, workers in the tech sector earn substantially more than workers in other industries with the same degree background. Second, graduates in computer science and electrical engineering—across all institution types—are far more likely to work in the tech sector than graduates of other fields or individuals without a postsecondary degree. Third, graduates from these fields (from any institution) earn significantly more when employed in tech relative to other tech workers.

The table also reports parental earnings, which highlight selection into technology-oriented careers. Across almost all degree categories, students from higher-income families are more likely to enter the tech sector. At the same time, the tech sector appears to promote mobility: for instance, graduates of computer science or electrical engineering degrees in colleges come from lower-income families than graduates of elite non-STEM programs, yet they earn more (23,600 NIS versus 19,500). Even outside the tech sector, their wages are slightly higher (13,500 versus 11,800), underscoring the central role of STEM graduates in the Israeli labor market.

Finally, although these figures reflect relatively young ages (30–32), the average salaries of computer science or electrical engineering graduates employed in tech place them among the very top of Israel's overall earnings distribution. For example, in our sample, graduates of non-elite universities with computer science or electrical engineering degrees who work in tech earn, on average, about 309,000 NIS (Table A.1) annually. For comparison, the 90th percentile of the national earnings distribution across all workers in 2018 was roughly 250,000 NIS (Danieli et al., 2024), which is approximately 280,000 in 2023 values. Using the

cohort-specific measures defined earlier, the dominance of STEM fields is also evident: more than 35 percent of computer science and electrical engineering graduates reach the top 5 percent of earnings in their test-taker cohort, compared with 8 percent in economics and smaller shares in other non-STEM fields.

Figure 2: Labor-Market Outcomes of Postsecondary Degree Graduates



Notes: The figure shows labor-market outcomes for all degree graduates in the sample (409,755 observations). The y-axis reports average monthly wages conditional on positive earnings (thousands of NIS), and the x-axis shows the share employed in the tech sector, both measured at ages 30–32. Circle size is proportional to the number of students in each program (field by institution type). Colors indicate institution type: grey for colleges, blue for non-elite universities, dark blue for elite universities. Labels are shown for programs with at least 3,000 students. Red horizontal and vertical lines mark lucrative STEM degrees, defined as those in the upper-right quadrant (above 17,500 NIS in monthly wages and at least 40 percent employed in tech). Students in the main analysis group (first-test scores 580–620) are excluded to avoid mechanical overlap with the treatment sample.

Recognizing these large returns, policymakers in Israel have sought to expand enrollment in selective STEM degrees. The Israeli Council for Higher Education recently launched a national initiative to boost participation in these programs.[10] This effort mirrors international policies promoting STEM enrollment (e.g., OECD, 2024). Understanding how young adults decide whether to apply to rewarding STEM degrees therefore has direct policy relevance for advanced economies facing persistent demand for high-skilled, tech-oriented workers.

This motivates our focus on selective STEM degrees. As the primary outcome, I define a selective STEM degree as computer science or electrical engineering at any university. I also use a broader empirical definition of lucrative STEM degrees: programs (fields by institution type) with average monthly wages above 17,500 NIS and at least a 40 percent tech-employment rate (see Table A.2). As shown in Figure 2, these two criteria coincide: all degrees with average wages above 17,500 NIS also have at least 40 percent tech employment. Within this group, almost all programs are offered by universities (elite or non-elite), with only two college programs included (computer science and a joint degree in computer science and electrical engineering).[11]

Based on these definitions, I construct indicator variables for applications (within five years of the first UPET attempt), enrollment, and degree attainment. Beyond STEM-focused indicators, I also track applications and enrollment in other fields, grouped by institution type and by field clusters following the Israeli Central Bureau of Statistics classification. Finally, I construct a "degree-associated wage" measure by assigning each individual the mean wage of graduates from their chosen program at ages 30–32. When multiple programs are relevant, I use the maximum value. This wage measure is divided into terciles: low (<12.5K NIS), mid (12.5–17.5K NIS), and high, which corresponds exactly to the lucrative STEM group (>17.5K NIS). Students in the main analysis group (those scoring 580–620 on their first test) are excluded when constructing this outcome to avoid mechanical overlap with the treatment sample.

2.3 Admissions to Selective STEM Degrees

Admission to selective STEM programs is highly competitive. While exact cutoffs are not recorded in the data, I observe enrolled students' admission-relevant scores. Among those

[10] Retrieved from https://che.org.il/hi-tech on June 23, 2025.

[11] Each definition has advantages. Focusing on the selective STEM dummy narrows attention to a small set of common, highly rewarding degrees, which simplifies interpretation and helps rule out alternative mechanisms such as admission advantages. Using the broader set allows quantifying the overall likelihood of pursuing high-return STEM degrees. We also test robustness to varying these definitions to ensure the results are not sensitive to specific cutoffs.

ever enrolled in selective STEM degrees, the median admission-relevant score is 685, and the 25th percentile is around 644.[12] These figures align with partial information on formal admission requirements in specific years (Appendix B.3). They show that a score just above 600 is rarely sufficient for admission unless paired with an exceptionally strong matriculation GPA. Thresholds for selective STEM are higher than for almost all other programs in Israel, with the exceptions of medicine (median 738; 25th percentile 721), which generally requires scores above 700.

The analysis therefore focuses on the 600-point threshold. Our quasi-experimental design exploits left-digit bias in test score perceptions for identification. While such bias could in principle arise at any round number, the 600 threshold is particularly informative for studying choices to pursue high-earning STEM degrees. Lower thresholds are too far below selective STEM admission cutoffs to matter, and the next higher round score (700) already grants direct access to most programs. Accordingly, the 600 threshold provides the most relevant setting for identifying the role of heuristic self-evaluation in STEM application decisions.[13]

Students scoring close to 600 on their first test are relatively high-ability—only about 30 percent of test-takers score above this mark—but still fall below the typical cutoff for selective STEM admission. The 600 threshold is therefore salient yet non-binding: students just above or just below 600 face essentially identical admission prospects, but the left-digit shift may alter how they interpret their performance. This setting also underpins the key identifying assumption: at the 600 cutoff, the only discontinuity is in perceived performance. Because admissions decisions are based strictly on objective criteria, left-digit bias does not directly affect admission chances. In Section 3, I provide empirical evidence that scoring just above 600 has no direct admissions advantage.

2.4 Study Sample

The main regression discontinuity analysis restricts attention to students scoring within 20 points of the 600 cutoff. It is useful to distinguish three subgroups of test-takers. First, I analyze Arab test-takers separately, given different background characteristics and attendance

[12]These values are much higher than for university students in any field (median 617; 25th percentile 526). Similar gaps appear if we use first scores (not necessarily admission-relevant): median 649 for selective STEM versus 574 for all university degrees. The differences are especially pronounced in quantitative scores (129 versus 113), compared with English (125 versus 112) or verbal (118 versus 110).

[13]Appendix C also documents effects of left-digit bias at the highest round number of 700. However, the smaller sample size makes it more difficult to examine potentially important heterogeneity. At lower round-number cutoffs, we also find meaningful effects of crossing them. Yet because these students are more likely to apply to colleges, and we lack full information on college applications, we restrict the analysis to the two highest round scores, where most students apply to universities.

at more segregated Arabic-language secondary schools. Among the remaining test-takers, age at first attempt provides a natural distinction. Most students in Israel begin university between ages 21 and 24, after completing compulsory military service (three years for men, two for women). About half take the test during high school or shortly thereafter (age $\leq$20), while the rest test later (Appendix Figure A.2a). Exemption from military service shifts Arab students' testing earlier and increases the window to retest before university entry. Accordingly, I analyze outcomes separately for three groups: younger ($\leq$20), older ($\geq$21), and Arab test-takers.

Table 1 reports summary statistics for the full sample and for each subgroup, both for all test-takers and for those within the regression discontinuity sample (581–619). Several patterns stand out. Arabs score substantially lower on their first test ($\approx$426 versus $\approx$553 for other young test-takers), which may help explain their lower enrollment in selective STEM degrees despite similar application rates. Comparing younger and older test-takers, initial scores are similar, but retesting is much more common for the younger group. Older students typically take the exam just before applying to university and thus have little time to retest and improve. Retesting overall is common: over half of test-takers retake the exam (Appendix Figure A.2b). Arabs' higher retesting rate likely reflects their lower initial scores.

For these reasons, the main analysis focuses on younger non-Arab test-takers. Arab students are underrepresented near the 600 cutoff because of large achievement gaps, limiting their role in the regression discontinuity analysis, while older students are less relevant given their limited scope for retesting and upward mobility in scores.

3 Empirical Strategy: Identification of Heuristic Self-Evaluation

This section outlines the empirical strategy for the key empirical goal of this study: to identify the role of heuristic-based self-evaluation in students' decisions and outcomes. Identification exploits the left-digit bias as a shock to test score perception. I implement a sharp regression discontinuity design, comparing students who score just above versus just below 600 on their first university entrance test. The underlying assumption is that potential outcomes are continuous at 600—so that the only discontinuity is in how students perceive their score.

Formally, let s_i denote the first score of test-taker i. Define the treatment as the perception boost due to left-digit change, $Z_i = \mathbf{1}\{s_i \geq 600\}$. For an outcome of interest Y_i, such as applying to selective STEM degrees, let $Y_i(j)$ denote the potential outcome if student i were assigned treatment status j, where $j \in \{0, 1\}$. In other words, $Y_i(1)$ is the outcome if the

Table 1: Summary Statistics by Group

	Main		Older		Arabs	
	(1)	(2)	(3)	(4)	(5)	(6)
RD Sample		V		V		V
Baseline Total Score	553.37	599.27	563.12	599.71	425.97	597.65
Quantitative Score	111.93	119.60	111.27	116.93	91.91	123.12
English Score	110.03	119.24	112.40	120.07	81.85	109.63
Qualitative Score	106.30	114.32	110.09	116.78	84.16	114.88
Max Total Score	588.38	635.93	573.82	610.41	463.99	656.61
Educated Parents Share	49.09	55.25	40.91	44.65	12.35	32.81
Female Share	63.48	60.33	52.87	50.32	61.13	50.94
Selective STEM Share	9.86	11.09	3.64	2.65	2.46	13.94
Tech Share	20.96	24.08	17.13	18.19	3.62	10.39
Wage (NIS thousands)	14.42	15.47	11.42	11.88	10.85	16.67
N	162548	18950	179436	21888	72396	1908

Note: This table presents summary statistics across different demographic groups. The data includes information on first test scores (total and by domain), maximum scores acheived by retesting, share of students with both parents having any postsecondary education, share of females, share of students in Selective STEM degrees (computer science or electrical engineering in universities), share of tech employment (at ages 30-32), monthly salaries (ages 30-32, measured in thousands of 2023 NIS). Values are rounded to two decimal places. Samples are defined by age at first test attempt: aged 20 or below, aged 21 or older, and a separate group of Arab test-takers. RD indicates whether the sample is restricted to those within the RD bandwidth of 20 points (focusing on below-cutoff individuals).

student receives a perception boost due to the scoring above 600, and $Y_i(0)$ is the outcome if they do not. The parameter of interest is the average treatment effect at the 600 threshold, that is, the expected difference between the potential outcomes with and without the boost, evaluated at the cutoff:

$$\tau \equiv \mathbb{E}[\, Y_i(1) - Y_i(0) \,|\, s_i = 600 \,]$$

Identification relies on the assumption that the conditional expectations of the potential outcomes are continuous at the cutoff. Formally, for $j \in \{0, 1\}$, $\mathbb{E}[Y_i(j) \mid s_i = s]$ is continuous at $s = 600$. The institutional context supports this assumption: scores are produced by a formal grading procedure, making manipulation implausible. I provide further empirical validation below. Under the continuity assumption, τ is identified and can be expressed as

$$\tau = \lim_{s \downarrow 600} \mathbb{E}[Y_i \mid s_i = s] - \lim_{s \uparrow 600} \mathbb{E}[Y_i \mid s_i = s] \tag{1}$$

Estimation of this discontinuity proceeds using a local linear regression, where the sample is restricted to a 20-point bandwidth ($s_i \in [581, 619]$), and I apply triangular kernel weights

centered at the cutoff. The coefficient of interest, τ, captures the effect of crossing 600 on the outcome in the first test. Standard errors are heteroskedasticity-robust, and clustered at the score level (Lee and Card, 2008). Robustness checks include bias-corrected estimation using the algorithm of Calonico et al. (2014) and honest inference following Kolesár and Rothe (2018). Data-driven bandwidth selectors generally yield similar bandwidths to the one employed in the main specification. I also examine robustness to alternative bandwidth choices (narrower and wider) and to specifications with lower- and higher-order polynomials.

Design validity requires that both the density of test scores and baseline covariates be smooth at 600. Figure A.1 presents the score distribution together with density discontinuity tests. A local linear density test (McCrary, 2008) detects a small, marginally significant discontinuity at 600 (-0.029 log points, s.e. 0.015, $p = 0.06$). Given the discrete nature of the running variable, this is likely a mechanical artifact rather than evidence of manipulation. Consistent with this interpretation, a parametric continuity check that exploits the (known) approximately normal distribution of scores and fits a quadratic within a ± 50-point window finds no significant discontinuity (-0.013 log points; s.e. 0.011; $p = 0.21$). This pattern is consistent with the institutional setting, where manipulation is implausible.

Table A.3 further supports validity by showing balance in baseline covariates using all tests in our sample. Panel A reports individual characteristics (age at test, female, born in Israel, religious school attendance, birth order). Panel B reports test timing and domain scores. Panel C reports family background (parental education, parental income at ages 14–16, number of siblings). Of 15 estimates, only two are significant at the 10 percent level, none at the 5 percent level, and all discontinuities are small in magnitude. Appendix Table A.4 shows analogous results for the main sample, with only one marginally significant discontinuity.

As an additional validity check, I regress the main outcome—application to selective university STEM programs within five years of the baseline test—on all predetermined variables (see Appendix B.1 for more details). Predicted values are smooth at 600 when using all tests and when focusing on the main sample (Figures A.3a and A.3b).[14] As a further falsification test, I examine discontinuities of the main outcome at non-round score cutoffs and show that the discontinuity at 600 is unique (Figures A.3c and A.3d).

Finally, I confirm that admission to selective STEM programs is stable around 600. Appendix Figure A.4 shows enrollment likelihood in university computer science or university electrical engineering degrees separately, as well as the joint likelihood of attending either program, around the 600 threshold. Unlike the main analysis, which relies on first scores, these figures

[14] Details on the prediction method appear in Appendix B.1; coefficients are reported in Table A.5.

use the admission-relevant scores (the highest scores achieved by each individual). This allows testing whether scoring above 600 provides any admission advantage.

When examining the likelihood of ever enrolling, the level is relatively low and does not increase discontinuously at 600. The slope does become steeper above 600, suggesting that scores just above 600 may start to matter for some students, likely depending on their Bagrut GPA. Restricting the analysis to applicants, the likelihood of enrollment is much higher, indicating that students who apply are already aware of their stronger chances relative to non-applicants. Yet here too, there is no discontinuous jump at 600. This reinforces the conclusion that scoring 600 does not provide a direct admission advantage. Appendix Figure A.5 replicates these patterns for the main sample.

4 Results: Impacts of Heuristic Self-Evaluation

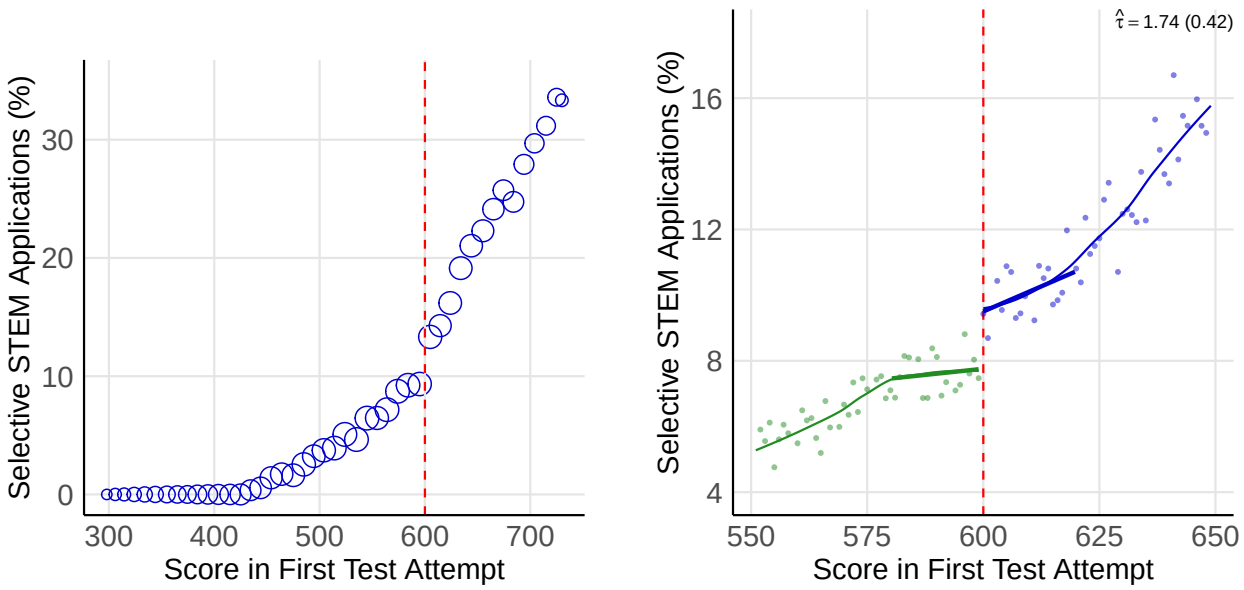
4.1 Applications to Selective STEM Degrees

Figure 3 shows that scoring just above 600 on the first attempt increases applications to selective STEM degrees. For all test-takers, the regression discontinuity estimate is a 1.7 percentage point increase (from 8.0 to 9.7 percent).[15] Panel (c) focuses on the main sample (age 20 or below), where the application rate rises by 3.8 percentage points from a 9.6 percent baseline. In contrast, Panel (d) shows no significant discontinuity for older test-takers, consistent with their lower baseline (5 percent) and limited scope to retest before applying.

Columns (2)–(6) of Table 2 expand the outcomes. The probability of applying to any lucrative STEM degree rises by 4.8 percentage points from a baseline of 19.0 percent. This increase is offset by a 2.7 percentage point decline in applications to lower-return programs, with no change in applications to degrees in the lowest associated wage tertile. As a result, the likelihood of submitting any application increases only modestly—by 2.2 percentage points from a baseline of 52.6 percent—and this estimate is noisy. The average wage associated with chosen programs increases by 740 NIS (≈200 USD) relative to a baseline of 16,400 NIS.

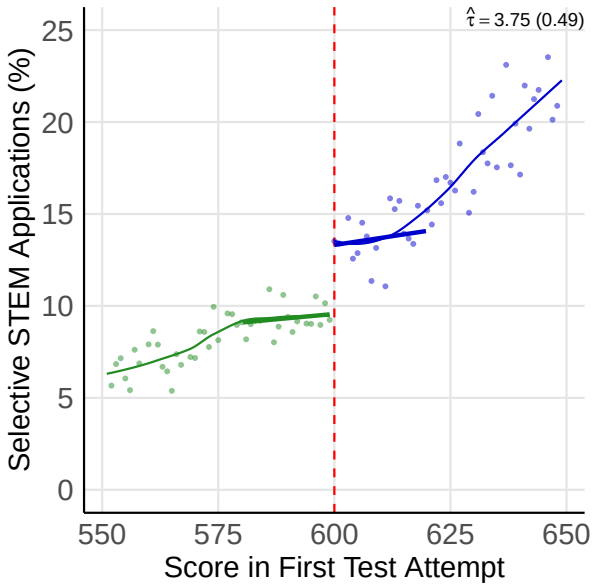
[15] Note that the figure also presents the estimated effect τ. Small differences from the tables may occur because it uses binned rather than raw data for estimation.

Figure 3: University Entrance Test Scores and Applications to Selective STEM Degrees

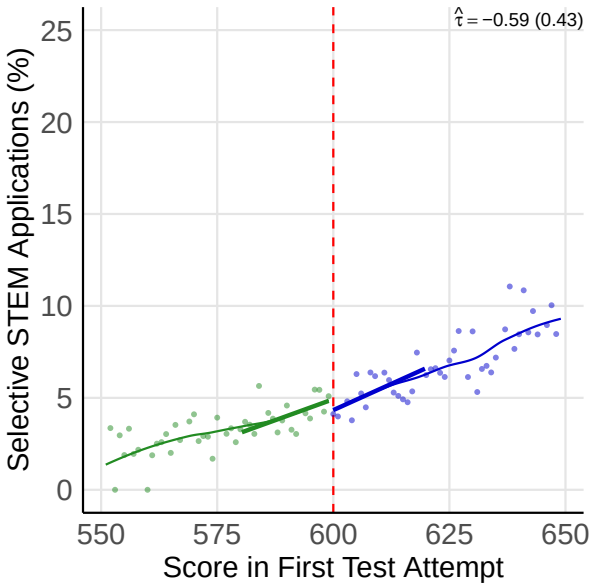


(a) Full Score Distribution

(b) Around the 600 Score Threshold



(c) Main (20 and Below)



(d) Older (21 and Above)

Notes: The figure shows that the probability of applying to selective STEM degrees increases sharply when students score just above 600 on their first university entrance test. The baseline sample includes 363,462 first-time test takers between 2000 and 2009. All panels plot the likelihood of applying to selective STEM degrees (y-axis) against initial scores (x-axis). Panel (a) displays mean application rates in 10-point score bins across the full range. Panel (b) focuses on 103,581 test takers within 50 points of the 600 threshold, plots raw means without binning, and overlays local linear regression fits (using a 20-point bandwidth) as well as a LOESS smoother estimated on the full window. The estimated coefficient τ from Equation 1, along with its standard errors, estimated using the binned data, is also reported. Panels (c) and (d) replicate the analysis for (c) the main sample (age at test 20 or below) and (d) the older sample (21 or above).. Red dashed vertical lines mark the 600 cutoff. Selective STEM degrees are defined as computer science and electrical engineering at universities. Outcomes are multiplied by 100.

Table 2: Effects of Crossing 600 on University Applications (Main Sample)

	(1)	(2)	(3)	(4)	(5)	(6)
	Selective STEM	Lucrative STEM	Other Mid	Other Low	Any	Associated Wage
	3.75***	4.76***	-2.72**	0.17	2.21	0.74***
	(0.49)	(0.78)	(1.38)	(1.16)	(1.62)	(0.18)
Mean	9.56	19.03	20.82	16.27	56.12	16.40
N	17,912	17,912	17,912	17,912	17,912	10,012

Notes: This table reports the estimated impact of crossing the 600-score cutoff on university applications. Columns (1)–(6) present estimates of τ from Equation 1, with robust standard errors in parentheses, clustered at the score level. The estimation sample includes first-time test takers in 2000–2009, aged 20 or below, with initial scores between 581 and 619. All outcomes are indicators for applications to universities within five years of the baseline test. Selective STEM includes computer science and electrical engineering; Lucrative STEM includes all degrees listed in Appendix Table A.2. Other Mid and Other Low classify remaining degrees by associated wage, as described in the text. Any refers to university degrees in any field. Associated wage is the average wage of students in the chosen application program (in thousands of NIS 2023 values), computed out of sample as described in the text. All indicator outcomes are multiplied by 100. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Results for other subsamples are reported in Table A.6. For older test-takers, most estimates show no significant changes, with only a marginally significant decline in the likelihood of applying to any university program. For Arab students, estimates are statistically imprecise because of the small sample size. This motivates the paper's focus on the main sample, with results for other subsamples reported in the appendix.

Panel A of Table A.7 reports additional results for the main sample. First, separating computer science and electrical engineering shows that the overall increase in selective STEM applications is driven mainly by computer science (3.15 percentage points), while the increase in electrical engineering is smaller (0.8 percentage points) and only marginally significant. Second, using an alternative definition of lucrative STEM degrees—all university programs with at least 40 percent tech employment—the likelihood of applying rises by five percentage points, very similar to the main definition of lucrative STEM (4.8). Finally, when extending the application horizon to ever applying, the increase in selective STEM applications remains similar (3.4 versus 3.8 percentage points) and highly statistically significant.

Heterogeneity of the effects. To examine heterogeneity in the impact of crossing the 600 threshold in the main sample, I estimate a local linear specification, which follows the same specification but interacts it with W_i, a personal characteristic of individual i. This is equivalent to estimating τ separately for each subsample, but allows inference on the difference in the estimated effect between different values of W_i. I validate that the results are robust to bias-corrected estimation following Calonico et al. (2025).

I examine three dimensions of heterogeneity: quantitative skills, socioeconomic status (SES), and gender. The first defines W_i as an indicator for a relative quantitative advantage on the UPET, equal to one if a student's quantitative score exceeds their scores in other domains by at least 5 points (about one quarter of the sample). In the main sample, most students do not apply to selective STEM degrees (baseline 10 percent), suggesting that these programs are not considered viable options for the majority. For students with stronger quantitative skills, however, selective STEM degrees appear more relevant. Table 3 shows that this group drives the discontinuity: their application rate rises by 8.4 percentage points from a baseline of 13.8 percent, whereas the increase for other students is only 2.3 percentage points. Results are similar when using the broader definition of lucrative STEM degrees.

Panel B of Table A.7 further shows that for quantitatively advantaged students, applications increase both in computer science (5.7 percentage points) and electrical engineering (3.2 percentage points). The results remain robust when using alternative outcome definitions, such as any degree with at least 40 percent tech employment, or when extending the horizon to ever applications. Panel C reports the results for the remaining students, where the effects are much smaller and become statistically insignificant once the horizon is extended to ever applications.

The second heterogeneity dimension is gender. Table 3 shows that the increase in STEM applications is larger for males (5.2 percentage points) than for females (2.5 percentage points), reflecting their different baseline application rates (15.5 percent versus 5.8 percent). The gap widens and becomes statistically significant when considering the broader set of lucrative STEM degrees: applications rise by 7.8 percentage points from a 25.8 percent baseline for males, compared with 2.3 percentage points from a 14.8 percent baseline for females.

The third dimension is socioeconomic status (SES), proxied by parental education. I define high SES as both parents having more than 12 years of schooling. Table 3 shows that the effects are slightly larger for high-SES students, but the differences are small and not statistically significant. Similar null results arise when SES is measured by parental income (above-median household earnings) or by combining both education and income.

Table 3: Heterogeneity of the Effects of Crossing 600 on University Applications (Main Sample)

	Difference	Z=0	Z=1
	(1)	(2)	(3)
	Z: Quantitative Advantage		
Selective STEM	6.05**	2.31***	8.36***
	(2.36)	(0.40)	(2.15)
		[8.16]	[13.75]
Lucrative STEM	7.38***	3.00***	10.38***
	(2.09)	(0.64)	(2.08)
		[16.73]	[25.89]
	Z: Female		
Selective STEM	-2.70	5.15***	2.45***
	(1.68)	(1.36)	(0.71)
		[15.54]	[5.83]
Lucrative STEM	-5.49**	7.80***	2.31*
	(2.20)	(1.66)	(1.18)
		[25.82]	[14.77]
	Z: High-SES		
Selective STEM	1.11	3.10**	4.21***
	(2.57)	(1.33)	(1.40)
		[8.64]	[10.31]
Lucrative STEM	3.79	2.58	6.38***
	(3.01)	(1.59)	(1.76)
		[18.03]	[19.84]

Notes: This table reports the estimated heterogeneity of the effects of crossing the 600-score cutoff. Column (1) reports estimates of the difference in estimated discontinuity effects across subgroups defined by variable Z. Columns (2)–(3) report the estimated impacts (τ from Equation 1) for subgroups with different values of Z, separately. Estimaion is based on a local linear regression, fully interacted with the indicator for Z. Robust standard errors, clustered at the score level, are in parentheses. Baseline means for each outcome and group are shown in square brackets. Outcomes are indicators for applying to a selective STEM degree (computer science and electrical engineering) or a lucrative STEM degree (fields listed in Table A.2). Heterogeneity dimensions include quantitative advantage (quantitative score at least five points above other domains), female indicator, and high-SES indicator (both parents with any postsecondary education). The estimation sample is the paper's main sample: first-time test takers in 2000–2009, aged 20 or younger, restricted to scores 581–619 on the first test. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Taken together, the heterogeneity results show that application responses are concentrated among students with quantitative readiness. All students in the sample are below the admission threshold. Crossing 600 provides a heuristic boost that nudges those already considering STEM to apply, with little effect on others. In this way, the effect reinforces pre-existing differences in STEM readiness.[16]

Alternative explanations. The evidence points to heuristic self-evaluation as the main channel. While I cannot observe self-perceptions directly, I test competing mechanisms. First, as discussed in Section 3, admission probabilities do not change discontinuously at 600: both scores just above and just below typically remain insufficient for entry into selective STEM programs. Hence, the effect cannot be explained by crossing a true admission threshold.

A natural alternative is that students mistakenly view 600 as an admission threshold. This implies that students scoring just above 600 would apply immediately, without further investment in improving their admission-relevant outcomes. Table A.8 contradicts this prediction. On average, retesting falls slightly (–2.4 percentage points from a 49.5 percent baseline), with no effect on maximum scores. Yet this average conceals sharp heterogeneity: students with a quantitative advantage retest more often (by 4.4 percentage points, though the estimate is noisy) and raise their maximum scores significantly (by 5.7 points), while others reduce retesting (–4.6 percentage points).

Linking retesting behavior directly to applications using interacted outcomes further supports the heuristic interpretation. Among quantitatively advantaged students, crossing 600 raises STEM applications by 8.4 percentage points, of which 5 points come from students who both apply and retake the exam. The probability of both applying to STEM and ultimately achieving a maximum score above 640 rises by 5.3 points from a 5.8 percent baseline. Thus, about 65 precent (5.3/8.4) of the marginal applicants not only apply to STEM but also substantially improve their scores before applying. This pattern suggests that the threshold motivates additional effort rather than complacency.

Additional evidence comes from matriculation (Bagrut) outcomes, another channel for raising admission scores. Columns (5)–(6) of Table A.8 show that students just above 600 complete on average 0.4 more matriculation credits (from a baseline of 29). Quantitatively advantaged students add 0.7 credits, mostly driven by additional credits in scientific subjects.

[16]The broader literature is mixed: some studies find that women and lower-SES students are especially sensitive to ability signals (Ahn et al., 2024; Graetz et al., 2023; Hakimov et al., 2023; Ugalde, 2022), whereas others show that correcting misperceptions has little effect on major choices (Zafar, 2013; Bestenbostel, 2020; Owen, 2022, 2023).

Together, these results rule out the hypothesis that students view 600 as sufficient for admission. Instead, they indicate that crossing 600 often motivates additional effort. Importantly, the heterogeneity suggests two distinct responses: for students with quantitative strengths, 600 acts as a validating signal that spurs ambition, while for others it provides a sense of closure and reduces effort. This duality is consistent with prior evidence on round-number heuristics in test score responses (Pope and Simonsohn, 2011; Goodman et al., 2020; Li and Qiu, 2023).

Finally, I examine social influence. Students' self-evaluations may be reinforced by peers or family members who share the same left-digit bias. While I cannot directly observe such influences, I study spillovers on younger siblings' decisions. Specifically, I analyze whether the closest-aged sibling (within five years) takes the test within three years after the baseline attempt (see more details in Appendix B.1). Table A.9 shows that younger siblings are 4.1 percentage points more likely to take the UPET within three years if an older sibling scored just above 600 (baseline 17.5 percent). The effect attenuates over longer horizons but remains positive. Interestingly, the effect is positive but imprecise in subsamples outside the main analysis and is not stronger among quantitatively advantaged students.

These findings suggest that social circles can reinforce heuristic self-evaluation. Crossing the threshold may raise students' own confidence and trigger encouragement from peers or family. This pattern relates to evidence on sibling spillovers in education (Joensen and Nielsen, 2018; Gurantz et al., 2020; Karbownik and Özek, 2023; Altmejd et al., 2021; Avdeev et al., 2024).

4.2 Degree Enrollment and Completion

Figure 4 shows that scoring just above 600 increases the probability of enrolling in a selective university STEM degree by 1.74 percentage points. Given that applications to these programs rise by 3.8 percentage points, this implies that 45 percent of the additional applicants ultimately enroll. Enrollment in any high-earning STEM program increases by 3.8 percentage points, indicating that students who do not enter selective STEM often still shift into other high-earning STEM fields. Additional estimates are shown in Panel A of Table 4. Overall, postsecondary enrollment is unaffected: nearly all students in this sample (96.3 percent) pursue any degree. Instead, the increase in STEM enrollment comes mostly at the expense of mid-range non-STEM programs, where enrollment declines by 3.5 percentage points.

Panel B of Table 4 reports results for degree completion.[17] Marginal applicants induced to apply by heuristic self-evaluation are no more likely to drop out than other students in the sample: Crossing 600 increases lucrative (selective) STEM graduation by 3 (1) percentage

[17]Table A.13 reports results for the subsamples of older and Arab test-takers, with no significant discontinuities.

point from a baseline of 16.7 (6.8) percent. Overall degree completion remains unchanged, with about 86.7 percent of students graduating. Figure A.6 shows no shift in degree enrollment timing. The rise in selective STEM enrollment occurs mainly 4–6 years after the baseline test, consistent with students retesting and improving their admission prospects over time.

Table 4: Effects of Crossing 600 on Postsecondary Degree Outcomes (Main Sample)

	(1)	(2)	(3)	(4)	(5)
A. Ever Enrollment					
	Selective STEM	Lucrative STEM	Other Mid	Other Low	Any
	1.74***	3.83***	-3.50***	-0.39	-0.06
	(0.53)	(0.77)	(1.30)	(1.59)	(0.28)
Mean	10.36	22.00	32.00	42.26	96.27
N	17,912	17,912	17,912	17,912	17,912
B. Completion					
	Selective STEM	Lucrative STEM	Other Mid	Other Low	Any
	0.98**	2.97***	-2.79**	-0.55	-0.37
	(0.41)	(0.84)	(1.33)	(1.60)	(0.72)
Mean	6.80	16.73	28.98	40.97	86.69
N	17,912	17,912	17,912	17,912	17,912

Notes: This table reports the estimated impact of crossing the 600-score cutoff on degree enrollment and completion. Columns (1)–(5) present estimates of τ from Equation 1, with robust standard errors in parentheses, clustered at the score level. The estimation sample includes first-time test takers in 2000–2009, aged 20 or below, with initial scores between 581 and 619. Panel A reports ever-enrollment outcomes, and Panel B reports degree completion. Selective STEM includes computer science and electrical engineering; Lucrative STEM includes all degrees listed in Appendix Table A.2. Other Mid and Other Low classify remaining degrees by associated salaries, as described in the text. Any refers to degrees in any field and institution type (universities or colleges). All indicator and rank outcomes are multiplied by 100. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

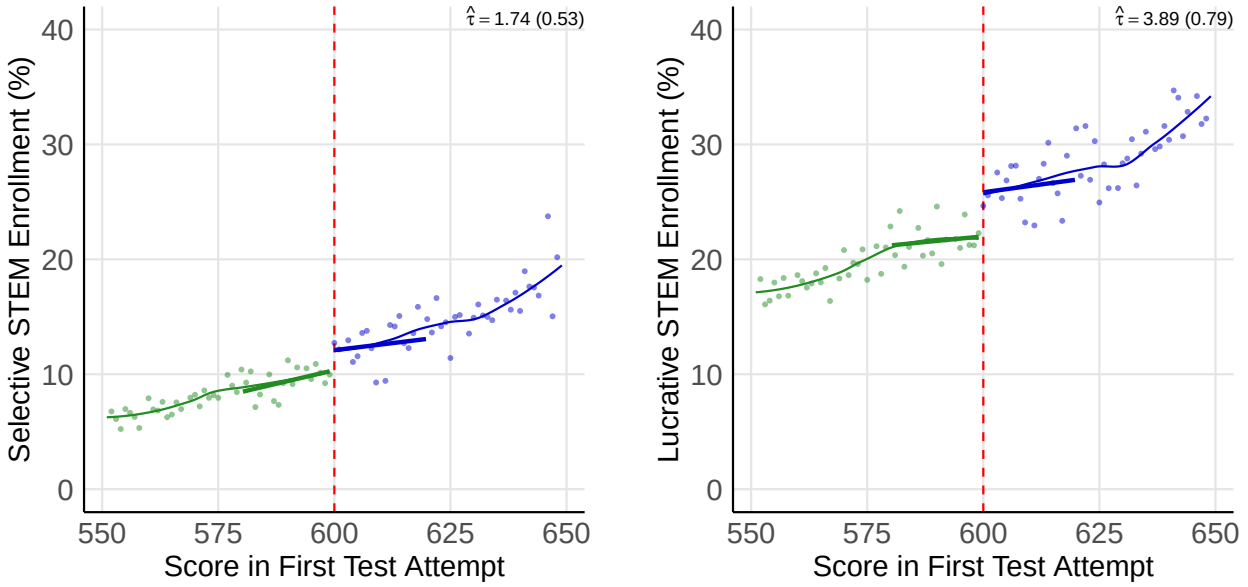
Additionally, I examine degree achievement using GPA data for graduates in the sample. Because grading standards vary across programs, I measure achievement as an indicator for graduating with a GPA above the program median (defined by field, institution type, and graduation year; see Appendix B.1). Table A.12 shows an insignificant 1.7 percentage-point

increase in the probability of graduating above the median for students just above 600. I also report average grades and find that students above 600 enroll in programs with slightly lower average grades, though the difference is insignificant.

The table also reports measures of pre-degree achievement ranking within programs. The likelihood of starting with a maximum UPET score below the program average decreases significantly, by 6.7 percentage points from a 24.5 percent baseline, for quantitatively advantaged students. Consistent with this, these students also attend programs with higher average UPET scores, by 9.4 points from a baseline of 631.3.

Overall, the results indicate that increased enrollment in demanding STEM programs did not reduce students' probability of graduating or cause them to receive very low grades. Instead, they perform better than would be predicted based on their UPET scores.

Figure 4: Effects of Crossing 600 on STEM Degree Enrollment (Main Sample)



(a) Selective STEM (%) (b) Lucrative STEM (%)

Notes: The figure shows the estimated effect of crossing the 600-score cutoff on STEM degree enrollment. The x-axis is the first university entrance test score. The y-axis indicates ever enrolling in a selective STEM degree (Panel a: computer science and electrical engineering in universities) or in a lucrative STEM degree (Panel b: fields listed in Table A.2). The sample includes 45,950 first-time test takers (age 20 or younger) in 2000–2009, and scored initially within 50 points of the 600 threshold Blue and green lines show local linear regression discontinuity fits(using a 20-point bandwidth) and a nonparametric (LOESS) fit estimated on the full window. The estimated coefficient τ from Equation 1, along with its standard errors, estimated using the binned data, is also reported. The estimation sample is restricted to 17,912 test takers who scored between 581 and 619. Outcomes are multiplied by 100.

Heterogeneity estimates are reported in Table A.10. Consistent with the application stage, the effects are concentrated among students with a quantitative advantage: ever enrolling in a lucrative STEM degree rises by 10.2 percentage points for this group, compared to only 1.9 points for others, the latter being only marginally significant. Table A.11 further shows that graduation from lucrative STEM degrees rises by 9.3 percentage points from a 24.8 percent baseline. Gender differences also mirror earlier results, with stronger effects for males (6.0 points) than for females (1.6 points), suggesting that the threshold reinforces pre-existing gender gaps in STEM careers. There is little difference in effects by socio-economic backgrounds.

Tables A.14 and A.15 examine effects by institution type, field, and degree level (undergraduate versus graduate) for both the main sample and the quantitatively advantaged subsample. Average university enrollment declines overall, but not for quantitatively advantaged students. For them, enrollment in elite universities shows an insignificant 2.5-point increase, driven by a marginally significant 2.6-point increase in enrollment in selective STEM degrees at these institutions. For this subsample, the wage associated with the attained degree rises by about 1,200 NIS (325 USD), and results are similar when averaging by field rather than by institution type. Among other common fields, enrollment declines in social sciences, paramedical sciences, and less-rewarding STEM (the latter estimate is noisy). I find no significant changes in enrollment in advanced or professional degrees (law and medicine).

4.3 Labor-Market Outcomes

Figure 5 and Table 5 present the estimated impacts of crossing the 600-point threshold on labor-market outcomes in 2022–2023. The results indicate substantial long-term benefits. Average monthly wage increases by 1,400 NIS, from a 24,200 NIS baseline, which is about 5.5 percent (4 log points).[18] Annual earnings rise by 13,200 NIS, about 6 percent (5 log points).

Crossing 600 also raises the likelihood of becoming a top earner. The average cohort-specific earnings rank increases by 3 percentile points. The probability of being in the top 10 percent rises by 3 percentage points from a baseline of 13 percent. Similarly, the probability of reaching the top 5 percent increases by 1.6 percentage points from a 6.9 percent baseline, and the probability of reaching the top 1 percent increases by 0.4 percentage points from a 0.5 percent baseline. Additionally, while overall employment rates (salaried or self-employed)

[18]Table A.18 reports results for the two other subgroups. Among Arab test-takers, earnings rise after crossing the threshold, but the increase is statistically insignificant. For older test-takers, by contrast, there is a significant decline in employment rate, likely reflecting their reduced retesting.

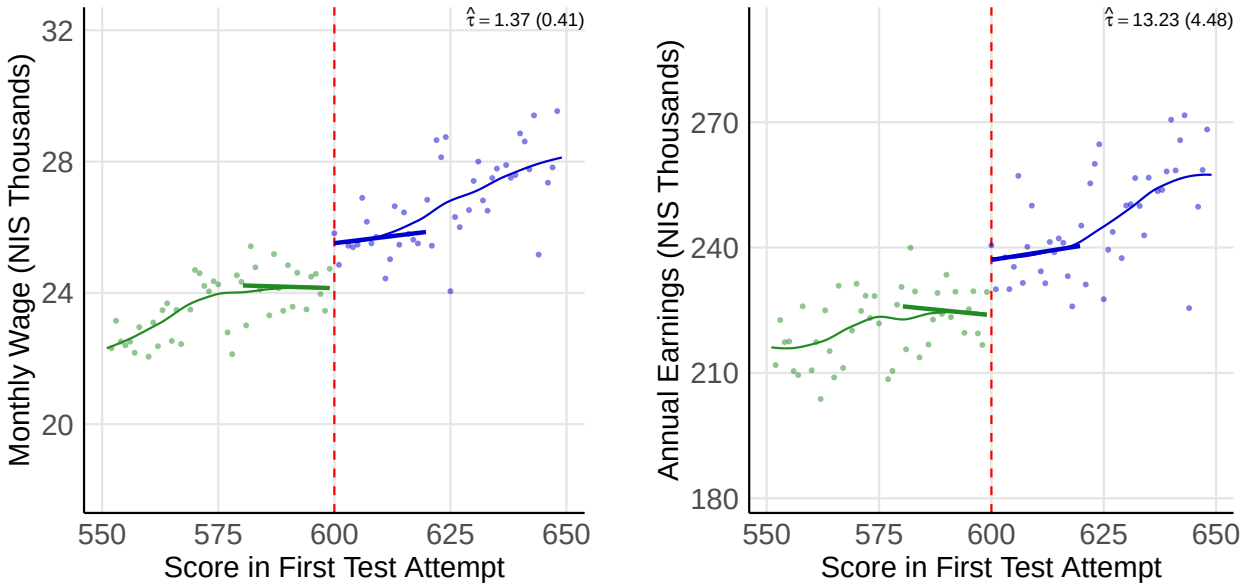
remain unchanged, the likelihood of working in the technology sector increases significantly, by 3.5 percentage points from a baseline of 26 percent.

Table 5: Effects of Crossing 600 on Labor-Market Outcomes in 2022-2023 (Main Sample)

	(1)	(2)	(3)	(4)
A. Earnings (NIS Thousands)				
	Annual Earnings	Log Annual	Monthly Wage	Log Monthly
	13.23***	5.15***	1.37***	4.13**
	(4.48)	(1.88)	(0.39)	(1.70)
Mean	223.81	1224.06	24.15	985.84
N	17,912	14,797	13,992	13,992
B. Rank and Top Earners				
	Rank	Top 10%	Top 5%	Top 1%
	2.97***	3.08***	1.59**	0.38*
	(0.82)	(0.65)	(0.78)	(0.21)
Mean	51.23	12.96	6.84	0.25
N	15,416	17,912	17,912	17,912
C. Employment				
	Any	Salaried	Self	Tech
	0.46	1.09	0.44	3.50***
	(0.92)	(0.95)	(0.89)	(0.89)
Mean	84.39	75.81	10.82	25.96
N	17,912	17,912	17,912	17,912

Notes: This table reports the estimated impact of crossing the 600 score cutoff on labor-market outcomes in 2022-2023. Columns (1)–(3) report the regression discontinuity estimates of τ from Equation 1, with robust standard errors in parentheses, clustered at the score level. The estimation sample is the paper's main sample: first-time test takers in 2000–2009, aged 20 or below, restricted to scores 581–619 on the first test. All earnings variables are measured in NIS thousands (2023 values) and other outcomes are measured as a percentage (multiplied by 100). Rank is the percentile rank of annual earnings among same-age test-takers in our sample. Top x is an indicator for being in the top x% of this distribution. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Figure 5: Effects of Crossing 600 on Earnings in 2022-2023 (Main Sample)



(a) Monthly Wage (b) Annual Earnings

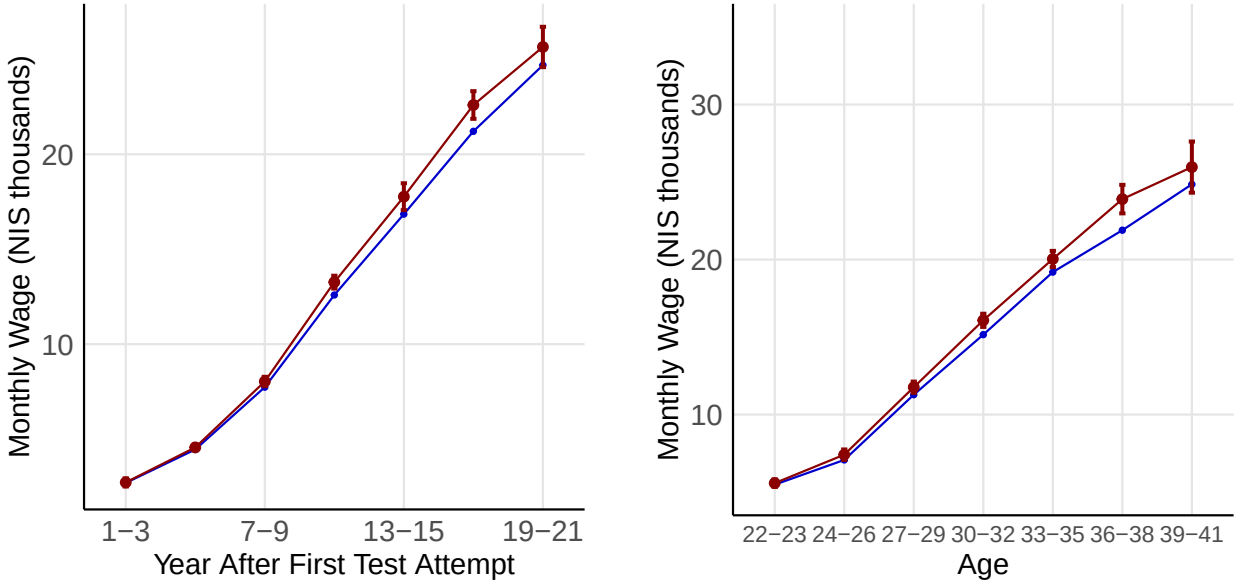
Notes: The figure shows the estimated effect of scoring just above 600 on labor-market outcomes in 2023. The x-axis is the first university entrance test score. The y-axis varies by panel: monthly wage conditional on any salaried income (a) and total annual earnings from all sources (b, including zeros), both in thousands of 2023 NIS. The sample includes 45,950 first-time test takers (age 20 or younger) in 2000–2009, and scored initially within 50 points of the 600 threshold Blue and green lines show local linear regression discontinuity fits(using a 20-point bandwidth) and a nonparametric (LOESS) fit estimated on the full window. The estimated coefficient τ from Equation 1, along with its standard errors, estimated using the binned data, is also reported. The estimation sample is restricted to 17,912 test takers who scored between 581 and 619.

Heterogeneity estimates in Table A.16 show that labor-market gains are concentrated among quantitatively advantaged students, consistent with their stronger responses at the application and enrollment stages. Their wages rise by 3,800 NIS, compared to an insignificant 600 NIS gain for others. Similar patterns appear by gender and SES: effects are larger for males and comparable across high- and low-SES students, though these estimates are imprecise. Table A.17 further shows that quantitatively advantaged students experience especially large returns, including higher probabilities of reaching the top of the earnings distribution.

These results show that the heuristic-induced shift in applications has persistent consequences: students are more likely to enroll in and complete high-earning STEM degrees, which in turn lead to careers in the tech sector, higher salaries, and entry into the very top of the earnings distribution. Thus, pursuing selective STEM degrees appears to be highly beneficial for the marginal applicants whose choices were shaped by the round-number

heuristic. Still, it is important to ask whether these gains could instead reflect other factors changing discontinuously at 600, such as employers' perceptions. Several patterns suggest that the observed labor-market effects stem from pursuing STEM degrees rather than from such factors. First, in line with standard STEM career trajectories, crossing 600 increases the likelihood of working in the tech sector and reaching the top of the earnings distribution pathways. Second, as shown in Figure 6, wage gains emerge only about a decade after the test, consistent with graduation and labor-market entry. If employers directly rewarded a first above-threshold score, one would expect immediate effects, which are not observed.[19] In addition, since students typically retake the UPET, the first score is unlikely to affect hiring decisions; if reported at all, it would be the highest score.

Figure 6: Effects of Crossing 600 on Monthly Wage Over Time (Main Sample)



(a) Period relative to test (b) Age

Notes: This figure shows baseline means (blue solid line) and implied means above the 600 cutoff (red solid line), where the latter equals the baseline mean plus the regression discontinuity estimate (τ from Equation 1), based on a local linear regression. The outcomes are monthly wage measured in three-year bins after the first test (a) or age in three-year bins (b). Error bars display 95 percent confidence intervals for the regression discontinuity estimate, based on robust standard errors clustered by score. The estimation sample is the paper's main sample: 17,912 first-time test takers in 2000–2009, aged 20 or younger, restricted to scores 581–619 on the first test.

[19] Indeed, most students work part-time during their studies—nearly 60 percent within three years after the test, rising to 76 percent within six years.

Importantly, Figure 6 shows that wage effects not only build gradually but also persist for up to 20 years after the test, through at least age 40, indicating a lasting career advantage rather than a temporary boost. Figure A.7 shows similar patterns for tech employment, suggesting that marginal STEM entrants remain in STEM careers at least until age 40. While prior work finds that returns to STEM decline over time (see, e.g., Deming and Noray, 2020), the marginal students in this setting continue to benefit at least through mid-career.

4.4 Robustness of the Estimation Results

Figure A.8 shows that the main estimates are stable across many specifications. I vary polynomial order (up to quadratic), use multiple bandwidths (larger windows for higher-order polynomials), and estimate models with and without the full set of controls listed in Table A.3. For each specification, I re-estimate the four key outcomes—applications to selective STEM, ever enrollment in lucrative STEM, monthly wages, and tech employment—and obtain remarkably similar effects.

I also estimate nonparametric MSE-optimal and honest regression discontinuity models following Calonico et al. (2015) and Kolesár and Rothe (2018). Results, shown in Table A.19, closely match the main analysis: the application effect is 4.0 percentage points in both nonparametric estimators versus 3.8 points in the baseline; wage gains are 1,400–1,600 NIS per month, similar to 1,400 in the baseline. The selected bandwidths are comparable (14–21), including the bandwidth of 20 used in the main analysis.

Additionally, I verify that the main heterogeneity pattern is robust using non-parametric estimation methods Calonico et al. (2025). Splitting students by quantitative advantage into three groups, the top group—identical to the subgroup emphasized throughout the paper—again drives the results (Table A.20). The estimated effects for this subsample are similar in magnitude. For example, applications increase by 10.8 percentage points (8.4 in the baseline) and wages by about 4,600 NIS (3,800 in the baseline).

Finally, I discuss the generalizability of these results. Appendix C examines left-digit bias at the 700 threshold, the highest round number in our setting. The sample is small and limited to top achievers (only 5 percent score above 700). Baseline STEM application rates are already high, and I do not detect a significant increase. Instead, I find a large rise in pursuing advanced degrees, including in STEM. Because UPET scores are not used in graduate admissions, this pattern may reflect a long-term self-evaluation effect. However, the small sample makes it difficult to study heterogeneity. Lower cutoffs are not analyzed

because students at those margins are more likely to apply to colleges, for which we lack complete application data.[20]

5 Returns to Choosing STEM

Crossing the 600 threshold induces uncertain students to pursue high-earning STEM programs and later achieve higher wages. I use this discontinuity in a fuzzy regression discontinuity design to estimate the returns to choosing STEM. Unlike admissions-margin studies (e.g., Bleemer and Mehta, 2022), my estimates capture application-margin effects, highlighting the returns for students who are induced to apply by a boost in self-evaluation. The results therefore inform policies that encourage uncertain applicants to raise aspirations and pursue demanding STEM degrees, rather than policies that simply expand program capacity.

5.1 Identification and Estimation

Recall from Section 3 that s_i denotes the first test score and the perception boost is $Z_i \equiv \mathbf{1}\{s_i \geq 600\}$, which serves as the instrument in this setting. Let $D_i^*(z) \in \{d_1, \ldots, d_K\}$ denote the potential degree program (defined by field and institution type) chosen by student i if $Z_i = z$, where K is the number of available programs. Let $Y_i(d^*)$ denote the potential outcome (e.g., wage at ages 30–32) if $D_i^* = d^*$.

I assume continuity. For each possible degree program d^* and each instrument value z, the conditional expectations $\mathbb{E}[Y_i(d^*) \mid s_i = s], \quad \Pr(D_i^*(z) = d^* \mid s_i = s)$ are continuous in s at 600. This assumption is supported by the institutional context and by the empirical evidence discussed earlier.

I further impose the exclusion restriction that the instrument affects outcomes only through degree choice, $Y_i(d^*, z) = Y_i(d^*)$ for all $d^* \in \{d_1, \ldots, d_K\},\ z \in \{0, 1\}$. Theoretically, the perception boost could affect labor-market outcomes directly—for example, by raising confidence and improving wage negotiations—even without altering degree choice. This channel is not unique to our setting and could, in principle, arise in contexts with admission-based cutoffs. However, the evidence does not support this mechanism. Labor-market effects appear only years after the test, when students graduate, rather than earlier as would be expected from confidence or negotiation effects, especially since most students in the sample work during and before their studies. Moreover, most students retake the test and improve their scores,

[20] When examining observed outcomes such as retesting and long-term enrollment, I find heterogeneous effects that likely reflect both channels described earlier. However, given the limited data on applications, such evidence is only suggestive.

making the initial score irrelevant in the long run. Finally, subgroups whose degree choices do not change at 600 show no subsequent wage gains.

The complication is that $D_i^*(z)$ is inherently a multivalued, unordered treatment. To fully identify degree-specific treatment effects, instrumental-variable approaches would require richer instruments—such as multiple admission cutoffs across programs—that generate more variation in treatment choice (e.g., Kirkeboen et al., 2016). In our case, such an approach is infeasible, since the analysis relies on a single binary instrument. To proceed, in the main analysis I focus on a simplification that treats degree choice as binary. This choice is motivated by the evidence: the primary margin of adjustment at the cutoff is a large increase in lucrative STEM enrollment, with no corresponding increase elsewhere. Before turning to the conditions required for binarization, I also note that an alternative transformation of the treatment can provide an interesting interpretation.

Returns relative to observational differences. I impose an ordering on programs based on their observed average wages, as described in Appendix D. Each program is assigned its mean wage, estimated outside of the regression discontinuity sample. Formally, $m(d) = \mathbb{E}[Y_i \mid D_i^* = d,\ s_i \notin [581, 619]]$. Then, each student is assigned the mean of their chosen program. This maps degree choice into an ordered continuous treatment. The resulting monotonicity condition is straightforward: crossing the 600 cutoff should weakly increase the probability of selecting higher-wage programs. I discuss this assumption and provide supporting evidence in the appendix.[21]

Under this monotonicity, together with continuity, exclusion, and relevance, the fuzzy regression discontinuity design identifies an interesting parameter: it compares the wage gains of compliers to the observational wage difference between their chosen and counterfactual programs. Appendix D shows this and the results. Estimates suggest that compliers' gains are not smaller, and likely larger, than the average returns of students in their chosen programs.

Returns to pursuing STEM. I next focus on the binary treatment of pursuing any lucrative STEM degree, defined as $D_i = \mathbf{1}\{D_i^* \in D^s\}$, where D^s is the set of degrees listed in Appendix A.2. This definition is motivated by the evidence: the main adjustment at the cutoff is a shift into lucrative STEM programs, not into other fields. The set D^s is determined by two conditions: high associated tech employment rates and high wages. Since all degrees with

[21] A similar argument appears in related empirical applications (Bleemer and Mehta, 2022; Bleemer, 2021). Such estimates are sometimes used in a non-heterogeneous framework to quantify the extent to which observational differences reflect causal returns, often referred to as the forecast coefficient (e.g., Angrist et al., 2024).

high wages also have high tech employment rates (see Figure 2), the definition is equivalent if based solely on wages. Hence, the binary treatment can be viewed as a binarization of the continuous treatment m, with $D_i = \mathbf{1}\{m(D_i^*) > 17.5\}$ (when m is measured in NIS thousands).

Define potential treatments $D_i(z)$ for $z \in \{0, 1\}$ as the degree choice under instrument value z. For an outcome Y_i (e.g., wage at ages 30–32), define potential outcomes $Y_i(d)$ for $d \in \{0, 1\}$ as the outcome under degree choice d. Observed variables satisfy $D_i = D_i(Z_i)$ and $Y_i = Y_i(D_i)$. The causal parameter of interest is the local average treatment effect (LATE) of pursuing a lucrative STEM degree for compliers at the cutoff:

$$\delta \equiv \mathbb{E}\big[\, Y_i(1) - Y_i(0) \,\big|\, s_i = 600,\ D_i(1) = 1,\ D_i(0) = 0\big]$$

Note that this parameter captures the wage gains from pursuing any lucrative STEM degree relative to the counterfactual field choices of compliers. In our setting, these counterfactuals consist mainly of non-STEM degrees in the social sciences and para-medical degrees (see Tables A.14 and A.15), which carry substantially lower descriptive wage premia in Israel.

This parameter is identified under the standard assumptions of a fuzzy regression discontinuity design. First, continuity follows from the multivalued condition discussed earlier. Second, I assume relevance: crossing the 600 cutoff is a relevant instrument for lucrative STEM enrollment, as shown by the discontinuity in applications and enrollments at the threshold. Third, I assume monotonicity, ruling out defiers. This is weaker than the monotonicity in m discussed earlier, but implied by it. Formally, $D_i(1) \geq D_i(0)$ for all i: students who would pursue a lucrative STEM degree when $Z_i = 0$ (below 600) would also do so when $Z_i = 1$ (at or above 600).

A possible violation could occur if some students retake the test only when scoring below 600, and a few of them then perform unusually well and enter STEM—students who would not have done so had they scored exactly 600. This scenario is unlikely, since the reduction in retaking is driven by students who are unlikely to consider STEM at all. Still, to address this concern, I also report instrumental variable estimates for quantitatively advantaged students, for whom such violations are implausible because retesting does not decline in this group.

Finally, I assume the exclusion restriction: crossing the 600 threshold influences wages only through enrollment in lucrative STEM degrees, $Y_i(d, z) = Y_i(d)$ for all z. This is the most demanding assumption in this setting, stronger than what was required earlier. There are two theoretical ways it could be violated. First, the perception shock could directly affect labor-market outcomes—for example, by boosting confidence and improving wage negotiations—without changing degree choice. As discussed above, the data do not support

such a channel. Second, the perception shock could shift students into different programs within the same treatment group D, such as moving among non-STEM degrees with different wage premia. This issue arises because we treat what is essentially a multivalued treatment as binary.

Earlier results suggest that neither violation drives the findings. The main change at the cutoff is a sharp increase in lucrative STEM enrollment, with no corresponding increase elsewhere, and the labor-market effects align with this mechanism: higher tech-sector employment and wage gains concentrated in the upper percentiles of the distribution. For this reason, I first present results under the exclusion restriction, then discuss possible violations more formally, and also report estimates that do not rely on this assumption.

Under all assumptions, the causal return parameter is identified by the fuzzy regression discontinuity estimand:

$$\delta = \frac{\lim_{s \downarrow 600} \mathbb{E}[Y_i \mid s_i = s] - \lim_{s \uparrow 600} \mathbb{E}[Y_i \mid s_i = s]}{\lim_{s \downarrow 600} \mathbb{E}[D_i \mid s_i = s] - \lim_{s \uparrow 600} \mathbb{E}[D_i \mid s_i = s]} \tag{2}$$

In the main specification, I estimate δ using local linear regression with a 20-point bandwidth around 600 and triangular kernel weights, consistent with the main approach throughout the paper. I also consider alternative specifications, including a kernel-weighted difference in means at the cutoff with a narrower 10-point bandwidth.

I also validate that the data do not strongly reject these assumptions. Within a 10-point window around the cutoff, I treat the instrument Z_i as as-if randomly assigned. In this setting, the fuzzy regression discontinuity reduces to a standard instrumental variable comparison of mean outcomes above and below the cutoff. Applying the joint instrument-validity test of Kitagawa (2015), the null of instrument validity cannot be rejected ($p = 0.40$). This provides evidence against violations of the key assumptions.

5.2 Main Estimation Results

Table 6 shows that in the main sample (Panel A), the estimated wage return in 2022–2023 is 26,200 NIS (≈7,000 USD), corresponding to about 80 log points. Both estimates are statistically significant, and the first-stage F-statistic is 17.1, which supports the relevance condition. I also follow the methodology of Abadie (2003) for describing outcomes for four groups defined by their compliance behavior. In 2022–2023, the mean wage of treated compliers (41,100 NIS) is slightly higher than that of always-takers (36,600 NIS), suggesting that compliers perform at least as well in STEM as other students in our sample. Similarly, the mean wage of untreated compliers (14,900 NIS) is somewhat lower than that of never-

takers (21,000 NIS), suggesting that compliers particularly benefit from STEM relative to non-STEM degrees, though the differences between these values are statistically insignificant.

Panel B shows results for the subsample of students with quantitative advantage, where violations of the underlying assumptions are less likely. The results are similar to those in Panel A, with estimates for the wage return of 30,000 NIS or about 100 log points in 2022-2023. Also, compliers appear to do slightly better, or at least no worse, in STEM relative to always-takers.

A similar pattern emerges when analyzing returns at different ages. The estimates are qualitatively stable across ages 30–32 and 33–35. For wages at ages 30–32, the return is 21,500 NIS (110 log points) in the main sample and 20,000 NIS (80 log points) in the quantitatively advantaged sample. At ages 33–35, the returns are 16,800 NIS (70 log points) in the main sample and 23,600 NIS (105 log points) in the quantitatively advantaged sample. Results are remarkably similar when estimated using a kernel-weighted difference in means at the cutoff (Table A.21). I also report results using total annual earnings as in Table A.22, where the magnitudes are broadly similar in the sample of quantitatively advantaged students, around 100 log points.

When considering the magnitude of these returns, it is useful to compare them to numbers in Table A.1. That table reports average wages for university graduates of selective STEM degrees who work in the tech sector: 27,000 NIS for non-elite universities and 30,000 NIS for elite universities. Note that the returns' estimates capture the broader group of lucrative STEM programs. Only about half of the compliers enter selective STEM, while the others pursue other lucrative STEM degrees, where the corresponding averages are 24,000 NIS and 27,000 NIS.

The estimated treated complier means at the same ages—30,100 NIS in the main sample and 30,500 NIS in the quantitatively advantaged sample—are therefore very much in line with these benchmarks, and perhaps slightly larger. This implies that marginal students in our setting achieve wages that are at least comparable to other graduates of lucrative STEM programs who successfully enter technology careers.

The estimated untreated complier means, 8,600 NIS in the main sample and 10,500 NIS in the quantitatively advantaged sample, are similar to the wages of non-tech workers at these ages. To complement this analysis, I re-estimate the models using employment in tech industries as the outcome. Results in Table A.23 show that most compliers transition from non-tech employment to tech-sector employment when treated.

Taken together, these results show that the heuristic perception boost shifts students into tech-oriented fields with high returns. Marginal STEM entrants graduate and work

Table 6: Returns to Lucrative STEM Degrees

	Never	Compliers		Always	
	E[Y(0)]	E[Y(0)]	E[Y(1)]	E[Y(1)]	**LATE**
	(1)	(2)	(3)	(4)	**(5)**
A. Sample: Main					
2022-2023 (NIS Thousands)	21.0	14.9	41.1	36.6	**26.15*** **
N=13,992, F=17.1	(0.2)	(6.4)	(5.6)	(0.5)	**(8.85)**
2022-2023 (Log)	9.7	9.7	10.5	10.3	**0.79** **
N=13,992, F=17.1	(0.0)	(0.2)	(0.2)	(0.0)	**(0.31)**
30-32 (NIS Thousands)	12.9	8.6	30.1	23.3	**21.54*** **
N=15,332, F=12.6	(0.1)	(5.5)	(3.1)	(0.3)	**(4.79)**
30-32 (Log)	9.3	9.1	10.2	9.8	**1.11** **
N=15,332, F=12.6	(0.0)	(0.4)	(0.2)	(0.0)	**(0.43)**
33-35 (NIS Thousands)	16.6	15.0	31.8	29.8	**16.76** **
N=13,860, F=16.1	(0.1)	(5.4)	(3.8)	(0.4)	**(6.24)**
33-35 (Log)	9.5	9.4	10.2	10.1	**0.71***
N=13,860, F=16.1	(0.0)	(0.3)	(0.2)	(0.0)	**(0.37)**
B. Sample: Quantitatively-Advantaged					
2022-2023 (NIS Thousands)	23.7	14.1	44.2	36.8	**30.10*** **
N= 3,664, F=20.4	(0.5)	(5.5)	(4.3)	(0.6)	**(7.89)**
2022-2023 (Log)	9.9	9.6	10.6	10.4	**1.02** **
N= 3,664, F=20.4	(0.0)	(0.3)	(0.2)	(0.0)	**(0.39)**
30-32 (NIS Thousands)	14.6	10.5	30.5	24.3	**20.02*** **
N= 3,986, F=15.3	(0.2)	(3.6)	(5.0)	(0.6)	**(4.96)**
30-32 (Log)	9.4	9.2	10.0	9.9	**0.79*** **
N= 3,986, F=15.3	(0.0)	(0.2)	(0.2)	(0.0)	**(0.26)**
33-35 (NIS Thousands)	19.3	10.3	33.9	30.9	**23.59*** **
N= 3,532, F=18.8	(0.4)	(5.2)	(3.5)	(0.6)	**(5.26)**
33-35 (Log)	9.7	9.2	10.2	10.2	**1.04*** **
N= 3,532, F=18.8	(0.0)	(0.3)	(0.2)	(0.0)	**(0.24)**

Notes: This table reports estimated returns to enrolling in lucrative STEM degrees (δ from Equation 2, estimated using local linear regression). It also reports estimated mean outcomes by type: never-takers, compliers (treated and untreated), and always-takers. Outcomes are monthly wages in thousands of NIS (2023 values), measured in 2022–2023 and at ages 30–32 and 33–35. The endogenous variable is an indicator for ever enrolling in a lucrative STEM degree (fields listed in Appendix Table A.2). The estimation sample in Panel A is the main sample: first-time test takers in 2000–2009, aged 20 or below, with scores between 581 and 619 on the first test. Panel B further restricts the sample to quantitatively advantaged students, in order to reduce concerns about the validity of the underlying assumptions (see discussion in Section 5). $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

in tech at least as often as their peers, do not have lower GPAs, and earn wages that are comparable or higher. By implication, untreated compliers—students just below 600 who do not pursue STEM—are likely undermatched in their alternative degree programs. Heuristic under-evaluation leads them to forgo substantial wage gains.

5.3 Relaxing the Exclusion Restriction

Given that degree choice is inherently multivalued, but the estimates above rely on a binary treatment, the exclusion restriction may appear overly strong. Appendix E formalizes the conditions under which such binarization is valid. The key requirement is that all compliers switch into or out of STEM, with no switching within the STEM group or within the non-STEM group. A visual diagnostic test in the spirit of Andresen and Huber (2021), reported in the appendix, supports the validity of the binarization.[22]

This finding can be rationalized as follows: if returns depend on ability mainly in STEM, then an ability boost will primarily shift students into these programs. While this is not a perfect description of reality, the restriction is not overly demanding in this context, since STEM degrees are both more challenging and more rewarding than the alternatives available to students in the sample.

Nevertheless, no diagnostic can fully rule out violations. To address this concern, I also consider the LATE under possible violations of the exclusion restriction due to within-group switches. The full discussion is presented in Appendix E, and summarized here. Switches within the STEM group—cases where $D_i(0) = D_i(1) = 1$ but $D_i^*(0) \neq D_i^*(1)$—are unlikely in our setting. As shown in Figure A.9, the share of students who pursue the very top-earning programs is extremely small, and compliers never reach them. This implies that if within-STEM switches occur, they must be between degrees with very similar wage premia.

The main concern is therefore switching within non-STEM fields, which could bias the binary returns estimates. For example, crossing the 600 threshold could shift students from sociology to economics—a transition that raises wages but leaves the STEM indicator unchanged. Such movements violate the exclusion restriction, since the instrument affects wages not only by changing STEM status D, but also by reallocating students across low-wage fields. This generates upward bias: the reduced form captures wage gains of both binary compliers and low-group movers, while the first stage reflects only the former.

[22] An alternative sufficient condition for binarization is the extensive-margin-only complier assumption, discussed in Rose and Shem-Tov (2024). However, they study a setting where all compliers move from no treatment to some treatment, which fits our context less well.

However, as mentioned earlier, I can separately estimate treated and untreated complier means following Abadie (2003). In particular, even if the exclusion restriction is violated by within–non-STEM switches, the treated complier mean $\mathbb{E}[Y(1) \mid C]$ remains point identified. By contrast, the untreated complier mean $\mathbb{E}[Y(0) \mid C]$ is not point identified in this case, and requires additional assumptions or bounds. Appendix E provides the formal discussion.

I focus on estimation using mean differences within a 10-point local window around the cutoff for simplicity. Within the subsample of quantitatively advantaged students at ages 30–32, the treated complier mean is 31,200 NIS (Table A.21). The counterfactual untreated-complier mean, however, may be biased by within-group switches. Still, it can be partially recovered from the outcome distribution of untreated individuals just below the cutoff, as shown in Figure A.10. By the fuzzy regression discontinuity assumptions, students below 600 who do not pursue STEM are a mixture of compliers (who would switch at 600) and never-takers. If compliers and never-takers have similar outcomes, the counterfactual mean is simply the group average, which is 14,200 NIS. This yields a LATE of 17,000 NIS—slightly smaller than the 20,000 NIS obtained in the main fuzzy regression discontinuity.

Since the share of compliers is point-identified from the first stage, and the mixture of compliers and never-takers below the cutoff is observed, we can construct weights for each group. A conservative lower bound is obtained by trimming the untreated distribution so that compliers are assumed to be those with the highest wages (similar to the approach of Lee, 2009). This yields worst-case bounds that still imply positive but insignificant returns of about 1,200 NIS.

Finally, I consider an alternative conservative counterfactual that trims by degree choice rather than by realized wages. Specifically, I restrict untreated individuals below the cutoff to those enrolled in programs with the highest observational wage premia. This approach relies on an auxiliary restriction detailed in the appendix, namely that counterfactual outcomes for compliers are weakly increasing in program-level premia $m(d)$. The resulting estimates are tighter but remain conservative, and they imply sizable and statistically significant gains of about 13,000 NIS.

6 Conclusion

This paper shows that uncertainty about ability and cognitive heuristics distort how students sort into higher education programs. Focusing on STEM—selective and the most financially rewarding degrees in Israel—I exploit a regression discontinuity at a round test score that carries no admissions advantage. Crossing this round number raises applications to selective

STEM programs by 3.8 percentage points from a baseline of 9.6 percent. Because the threshold typically sits below admission cutoffs and provides no formal advantage, the effect reflects left-digit bias, in which students interpret a round number as a meaningful signal of ability.

Following students for over two decades reveals that these choices shaped by heuristics translate into large and persistent gains. Treated students improve their scores, enroll in and complete high-return STEM programs, and ultimately earn substantially more. Monthly wages rise by about 1,400 NIS (4 log points), annual earnings by 13,200 NIS (5 log points), and employment in the technology sector increases by 3.5 percentage points. Cohort earnings rank shifts upward by three percentiles, with higher probabilities of reaching the top of the distribution.

An instrumental variables analysis that uses the round-score shock to instrument STEM choice shows very large returns, with wages approximately doubling. While these estimates rely on a non-trivial exclusion restriction, I provide supporting evidence, and even bounds that relax exclusion still imply sizeable gains. The relevant policy margin here differs from the usual admission margin, which speaks mainly to the benefits of expanding seats. My estimates capture the application margin, highlighting an additional tool: encouraging capable yet uncertain students to apply. However, the results also indicate that such interventions are unlikely to close enrollment gaps. Belief boosts operate only where quantitative preparation is already in place, so reducing gender or socioeconomic disparities requires earlier pipeline investments.

Limitations are important to note. First, I do not observe direct measures of self-evaluation or beliefs. While the analysis rules out several alternative channels and the evidence points to self-evaluation as the most plausible driver, the mechanism cannot be directly verified. External validity may also be limited by the identification strategy. To provide clean causal inference in administrative data, I focus on a specific round-number threshold, where uncertain students receive a quasi-random self-evaluation boost from a salient heuristic. This design isolates a narrow margin, but self-evaluation likely operates more broadly in students' choices, and the effects documented here are best interpreted as a local manifestation of a general phenomenon. Future work could strengthen external validity by linking administrative records to survey or experimental measures of self-beliefs, probing the underlying psychological mechanisms more directly, and examining effects across the score distribution. Finally, the large estimated STEM returns for lower-testing students here are most relevant in economies with sustained demand for STEM workers. They may be weaker elsewhere, though many advanced economies share this feature, and the results are consistent with evidence from

other contexts showing that less-prepared students who enter STEM still experience sizable gains.

This paper illustrates the powerful role of uncertainty and heuristics in shaping human capital investment and long-term career outcomes. Findings expose a key weakness of modern postsecondary admission systems, which rely heavily on test-based meritocracies where students compete for higher scores to access selective and rewarding degrees. The deep uncertainty young adults face, combined with their reliance on heuristics, can cause equally qualified students to take divergent educational and career paths. This underscores the need for transparent communication and guidance to help students interpret signals of ability and make informed choices.

References

Abadie, Alberto, "Semiparametric instrumental variable estimation of treatment response models," *Journal of econometrics*, 2003, *113* (2), 231–263.

Achdut, Leah, Elad Gutman, Idan Lipiner, Inbal Maayan, and Noam Zussman, "The Wage Premium on Higher Education: Universities and Colleges," *Israel Economic Review*, 2019, *17* (2), 29–99.

Ahn, Tom, Peter Arcidiacono, Amy Hopson, and James R. Thomas, "Equilibrium Grading Policies With Implications for Female Interest in STEM Courses," *Econometrica*, 2024, *92* (3), 849–880.

Altmejd, Adam, Andrés Barrios-Fernández, Marin Drlje, Joshua Goodman, Michael Hurwitz, Dejan Kovac, Christine Mulhern, Christopher Neilson, and Jonathan Smith, "O brother, where start thou? Sibling spillovers on college and major choice in four countries," *The Quarterly Journal of Economics*, 2021, *136* (3), 1831–1886.

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

_ , Peter Arcidiacono, and Arnaud Maurel, "The analysis of field choice in college and graduate school: Determinants and wage effects," in "Handbook of the Economics of Education," Vol. 5, Elsevier, 2016, pp. 305–396.

Andresen, Martin E and Martin Huber, "Instrument-based estimation with binarised treatments: issues and tests for the exclusion restriction," *The Econometrics Journal*, 2021, *24* (3), 536–558.

Andrews, Rodney J, Scott A Imberman, Michael F Lovenheim, and Kevin Stange, "The returns to college major choice: Average and distributional effects, career trajectories, and earnings variability," *Review of Economics and Statistics*, 2024, *106* (6), 1428–1472.

Angrist, Joshua, Peter Hull, Parag A. Pathak, and Christopher R. Walters, "Credible School Value-Added with Undersubscribed School Lotteries," *Review of Economics and Statistics*, 2024, *106* (1).

Arcidiacono, Peter, Esteban Aucejo, Arnaud Maurel, and Tyler Ransom, "College attrition and the dynamics of information revelation," *Journal of Political Economy*, 2025, *133* (1), 53–110.

_ , V Joseph Hotz, and Songman Kang, "Modeling college major choices using elicited measures of expectations and counterfactuals," *Journal of Econometrics*, 2012, *166* (1), 3–16.

Avdeev, Stanislav, Nadine Ketel, Hessel Oosterbeek, and Bas van der Klaauw, "Spillovers in Fields of Study: Siblings, Cousins, and Neighbors," *Journal of Public Economics*, 2024, *238*, 105193.

Avery, Christopher, Oded Gurantz, Michael Hurwitz, and Jonathan Smith, "Shifting college majors in response to advanced placement exam scores," *Journal of Human Resources*, 2018, *53* (4), 918–956.

Bestenbostel, Adam, "Do Grade Signals Drive the Gender Gap in STEM? Evidence from a Regression Discontinuity," in "2020 APPAM Fall Research Conference" APPAM 2020.

Black, Sandra E., Jeffrey T. Denning, and Jesse Rothstein, "Winners and losers? The effect of gaining and losing access to selective colleges on education and labor market outcomes," *American Economic Journal: Applied Economics*, 2023, *15* (1), 26–67.

Bleemer, Zachary, "Top Percent Policies and the Return to Postsecondary Selectivity. Research & Occasional Paper Series: CSHE. 1.2021.," *Center for Studies in Higher Education*, 2021.

_ , "Affirmative action, mismatch, and economic mobility after California's Proposition 209," *The Quarterly Journal of Economics*, 2022, *137* (1), 115–160.

_ **and Aashish Mehta**, "Will studying economics make you rich? A regression discontinuity analysis of the returns to college major," *American Economic Journal: Applied Economics*, 2022, *14* (2), 1–22.

Calonico, Sebastian, Matias D Cattaneo, and Rocio Titiunik, "Robust nonparametric confidence intervals for regression-discontinuity designs," *Econometrica*, 2014, *82* (6), 2295–2326.

_ **,** _ **, and** _ , "Optimal data-driven regression discontinuity plots," *Journal of the American Statistical Association*, 2015, *110* (512), 1753–1769.

_ **,** _ **, Max H Farrell, Filippo Palomba, and Rocio Titiunik**, "Treatment effect heterogeneity in regression discontinuity designs," *arXiv preprint arXiv:2503.13696*, 2025.

Daly, Moira, Mathias Fjællegaard Jensen, and Daniel le Maire, "University admission and the similarity of fields of study: Effects on earnings and skill usage," *Labour Economics*, 2022, *75*, 102118.

Danieli, Oren, Shai Gilat, and Dor Leventer, "Top Income Inequality in Israel," *Israel Economic Quarterly*, 2024, *22* (1), 157–244.

Deming, David J and Kadeem Noray, "Earnings dynamics, changing job skills, and STEM careers," *The Quarterly Journal of Economics*, 2020, *135* (4), 1965–2005.

Denning, Jeffrey T, Richard Murphy, and Felix Weinhardt, "Class rank and long-run outcomes," *Review of Economics and Statistics*, 2023, *105* (6), 1426–1441.

Diamond, Rebecca and Petra Persson, "The Long-Term Consequences of Teacher Discretion in Grading of High-Stakes Tests," Technical Report w22207, National Bureau of Economic Research 2016.

Dube, Arindrajit, Alan Manning, and Suresh Naidu, "Monopsony and employer misoptimization explain why wages bunch at round numbers," *American Economic Review*, 2025, *115* (8), 2689–2721.

Dynarski, Susan, Carmello J Libassi, Katherine Michelmore, and Stephanie Owen, "Closing the gap: The effect of reducing complexity and uncertainty in college pricing on the choices of low-income students," *American Economic Review*, 2021, *111* (6), 1721–1756.

Ebenstein, Avraham, Victor Lavy, and Sefi Roth, "The long-run economic consequences of high-stakes examinations: Evidence from transitory variation in pollution," *American Economic Journal: Applied Economics*, 2016, *8* (4), 36–65.

Elsner, Benjamin and Ingo E Isphording, "A big fish in a small pond: Ability rank and human capital investment," *Journal of Labor Economics*, 2017, *35* (3), 787–828.

_ , _ , and Ulf Zölitz, "Achievement rank affects performance and major choices in college," *The Economic Journal*, 2021, *131* (640), 3182–3206.

Goodman, Joshua, Oded Gurantz, and Jonathan Smith, "Take two! SAT retaking and college enrollment gaps," *American Economic Journal: Economic Policy*, 2020, *12* (2), 115–158.

Graetz, Georg, Björn Öckert, and Oskar Nordström Skans, "Parental Education and the Responses to Higher SAT Scores," *Journal of Human Resources*, 2023. Advance online publication.

Gurantz, Oded, Michael Hurwitz, and Jonathan Smith, "Sibling effects on high school exam taking and performance," *Journal of Economic Behavior & Organization*, 2020, *178*, 534–549.

Hakimov, Rustamdjan, Renke Schmacker, and Camille Terrier, "Confidence and college applications: Evidence from a randomized intervention," Technical Report, Working Paper 2023.

Hastings, Justine S, Christopher A Neilson, and Seth D Zimmerman, "Are some degrees worth more than others? Evidence from college admission cutoffs in Chile," Technical Report, National Bureau of Economic Research 2013.

Heinesen, Eskil, Christian Hvid, Lars Johannessen Kirkebøen, Edwin Leuven, and Magne Mogstad, "Instrumental Variables with Unordered Treatments: Theory and Evidence from Returns to Fields of Study," *Journal of Labor Economics*, 2024, *42* (4).

Joensen, Juanna Schrøter and Helena Skyt Nielsen, "Spillovers in education choice," *Journal of Public Economics*, 2018, *157*, 158–183.

Kapor, Adam, "Transparency and percent plans," Technical Report w32372, National Bureau of Economic Research 2024.

Karbownik, Krzysztof and Umut Özek, "Setting a Good Example? Examining Sibling Spillovers in Educational Achievement Using a Regression Discontinuity Design," *Journal of Human Resources*, 2023, *58* (5), 1567–1607.

Kirkeboen, Lars J, Edwin Leuven, and Magne Mogstad, "Field of study, earnings, and self-selection," *The Quarterly Journal of Economics*, 2016, *131* (3), 1057–1111.

Kitagawa, Toru, "A test for instrument validity," *Econometrica*, 2015, *83* (5), 2043–2063.

Kolesár, Michal and Christoph Rothe, "Inference in regression discontinuity designs with a discrete running variable," *American Economic Review*, 2018, *108* (8), 2277–2304.

Krill, Ze'ev, Assaf Geva, and Tslil Aloni, "The effect of the field of study on the higher education wage premium—Evidence from Israel," *Labour*, 2019, *33* (3), 388–423.

Lacetera, Nicola, Devin G. Pope, and Justin R. Sydnor, "Heuristic Thinking and Limited Attention in the Car Market," *American Economic Review*, 2012, *102* (5), 2206–2236.

Landaud, Fanny, Éric Maurin, Barton Willage, and Alexander Willén, "The Value of a High School GPA," *Review of Economics and Statistics*, 2024, pp. 1–24.

Lavy, Victor and Rigissa Megalokonomou, "Persistency in Teachers' Grading Bias and Effects on Longer-Term Outcomes: University Admissions Exams and Choice of Field of Study," Technical Report 26021, National Bureau of Economic Research 2019.

_ **and Yoav Goldstein**, "Gifted Children Programs' Short and Long-Term Impact: Higher Education, Earnings, and the Knowledge-Economy," Technical Report w29779, National Bureau of Economic Research 2022.

Lee, David S., "Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects," *The Review of Economic Studies*, 2009, *76* (3), 1071–1102.

Lee, David S and David Card, "Regression discontinuity inference with specification error," *Journal of Econometrics*, 2008, *142* (2), 655–674.

Li, Hongbin and Xinyao Qiu, "Heuristics in Self-Evaluation: Evidence from the Centralized College Admission System in China," *Review of Economics and Statistics*, 2023.

Li, Hongyan and Xing Xia, "Grades as Signals of Comparative Advantage: How Letter Grades Affect Major Choices," *Journal of Economic Behavior & Organization*, 2024, *221*, 104891.

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

Murphy, Richard and Felix Weinhardt, "Top of the class: The importance of ordinal rank," *The Review of Economic Studies*, 2020, *87* (6), 2777–2826.

Ng, Kevin and Evan Riehl, "The returns to STEM programs for less-prepared students," *American Economic Journal: Economic Policy*, 2024, *16* (2), 37–77.

OECD, *Education at a Glance 2024: OECD Indicators*, Paris: OECD Publishing, 2024.

Olenski, Andrew R., André Zimerman, Stephen Coussens, and Anupam B. Jena, "Behavioral Heuristics in Coronary-Artery Bypass Graft Surgery," *New England Journal of Medicine*, 2020, *382* (8), 778–779.

Owen, S, "Ahead of the curve: Grade signals, gender, and college major choice," Technical Report, Working paper 2022.

Owen, Stephanie, "College Major Choice and Beliefs About Relative Performance: An Experimental Intervention to Understand Gender Gaps in STEM," *Economics of Education Review*, 2023, *97*, 102479.

Papay, John P, Richard J Murnane, and John B Willett, "The impact of test score labels on human-capital investment decisions," *Journal of Human Resources*, 2016, *51* (2), 357–388.

Pope, Devin and Uri Simonsohn, "Round numbers as goals: Evidence from baseball, SAT takers, and the lab," *Psychological science*, 2011, *22* (1), 71–79.

Rose, Evan K and Yotam Shem-Tov, "On Recoding Ordered Treatments as Binary Indicators," *Review of Economics and Statistics*, 2024.

Shurtz, Ity, "Heuristic thinking in the workplace: Evidence from primary care," *Health Economics*, 2022, *31* (8), 1713–1729.

Smith, Jonathan, Michael Hurwitz, and Christopher Avery, "Giving college credit where it is due: Advanced Placement exam scores and college outcomes," *Journal of Labor Economics*, 2017, *35* (1), 67–147.

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

Stinebrickner, Todd and Ralph Stinebrickner, "Learning about academic ability and the college dropout decision," *Journal of Labor Economics*, 2012, *30* (4), 707–748.

Strulov-Shlain, Avner, "More Than a Penny's Worth: Left-Digit Bias and Firm Pricing," *Review of Economic Studies*, 2023, *90* (5), 2612–2645.

Ugalde, Maria Paola, "Gender, grade sensitivity, and major choice," 2022.

Wiswall, Matthew and Basit Zafar, "Determinants of college major choice: Identification using an information experiment," *The Review of Economic Studies*, 2015, *82* (2), 791–824.

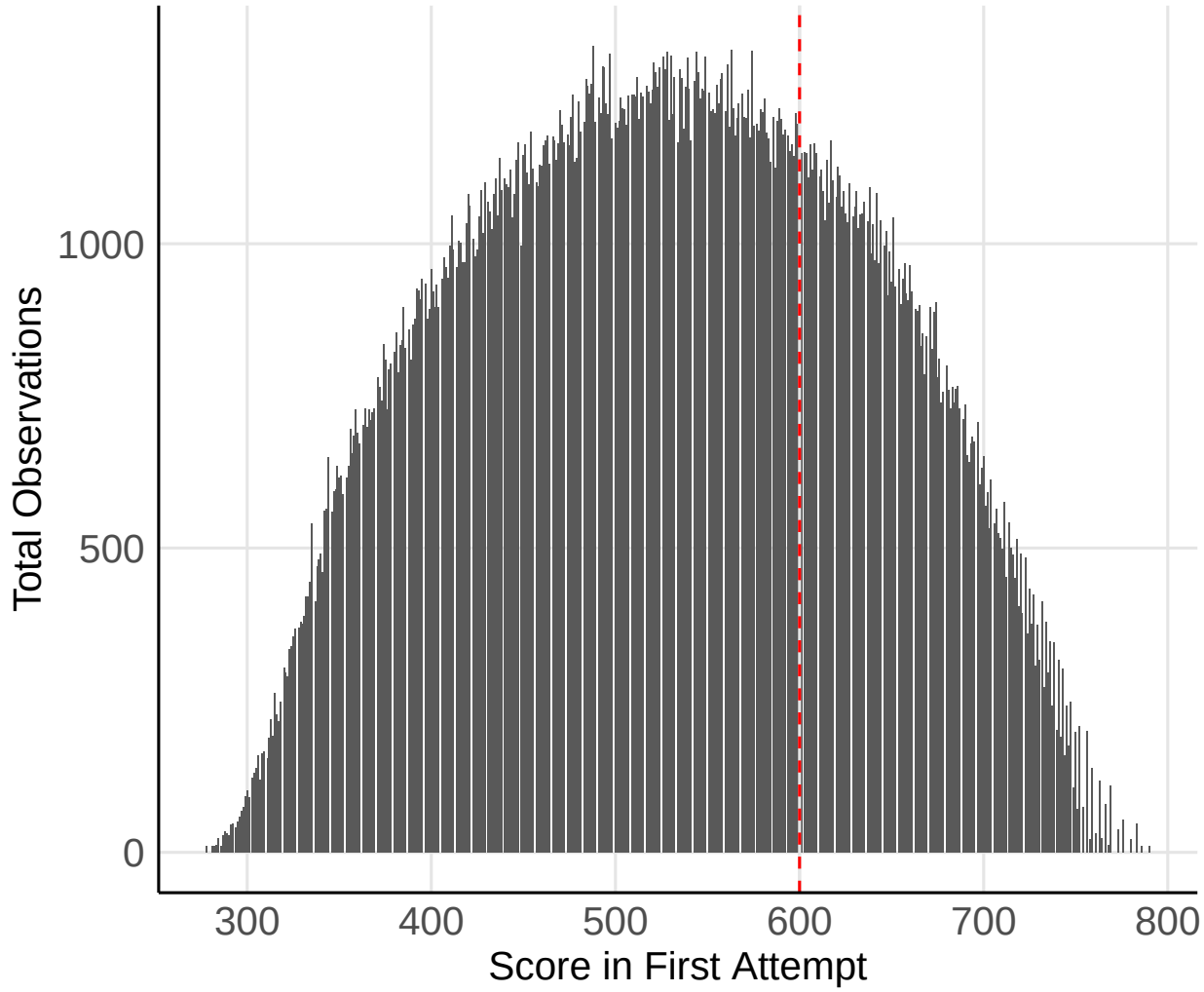
Zafar, Basit, "How do college students form expectations?," *Journal of Labor Economics*, 2011, *29* (2), 301–348.

_ , "College major choice and the gender gap," *Journal of Human Resources*, 2013, *48* (3), 545–595.

For Online Publication

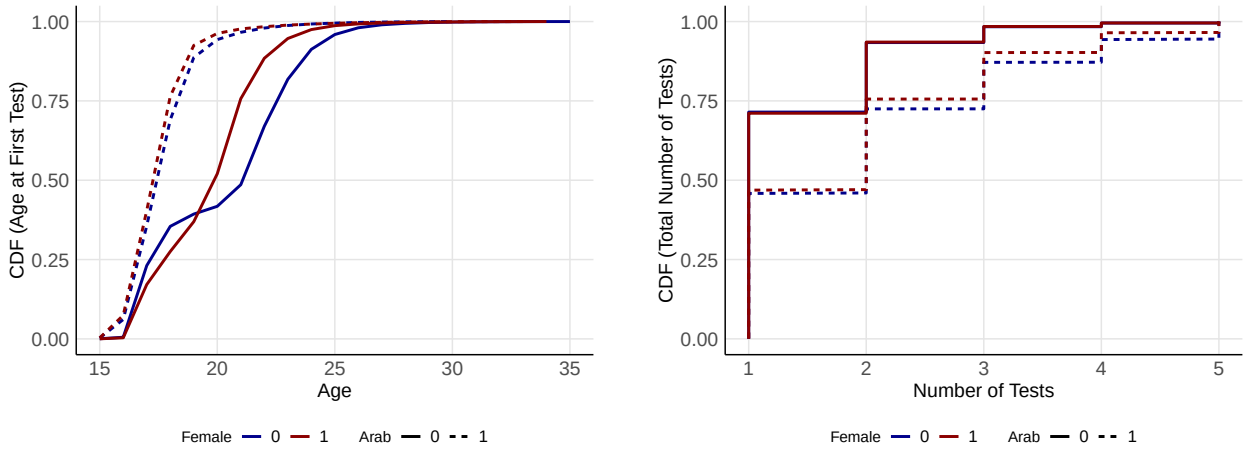
Appendix A Additional Figures and Tables

Figure A.1: University Entrance Test Score Distribution, First Attempts



Notes: The figure displays the distribution of first-attempt university entrance test scores for 359,122 individuals who took their first test between 2000 and 2009. The red dashed line indicates the 600 score threshold. A local linear density test (McCrary, 2008) detects a small, marginally significant discontinuity at 600 (-0.029 log points; s.e. 0.015; $p = 0.06$). Given the discrete nature of the running variable, this is likely a mechanical artifact rather than evidence of manipulation. Consistent with this interpretation, a parametric continuity check that exploits the approximately normal distribution of scores and fits a quadratic within a ± 50-point window finds no significant discontinuity (-0.013 log points; s.e. 0.011; $p = 0.21$). This pattern is consistent with the institutional setting, where manipulation is implausible.

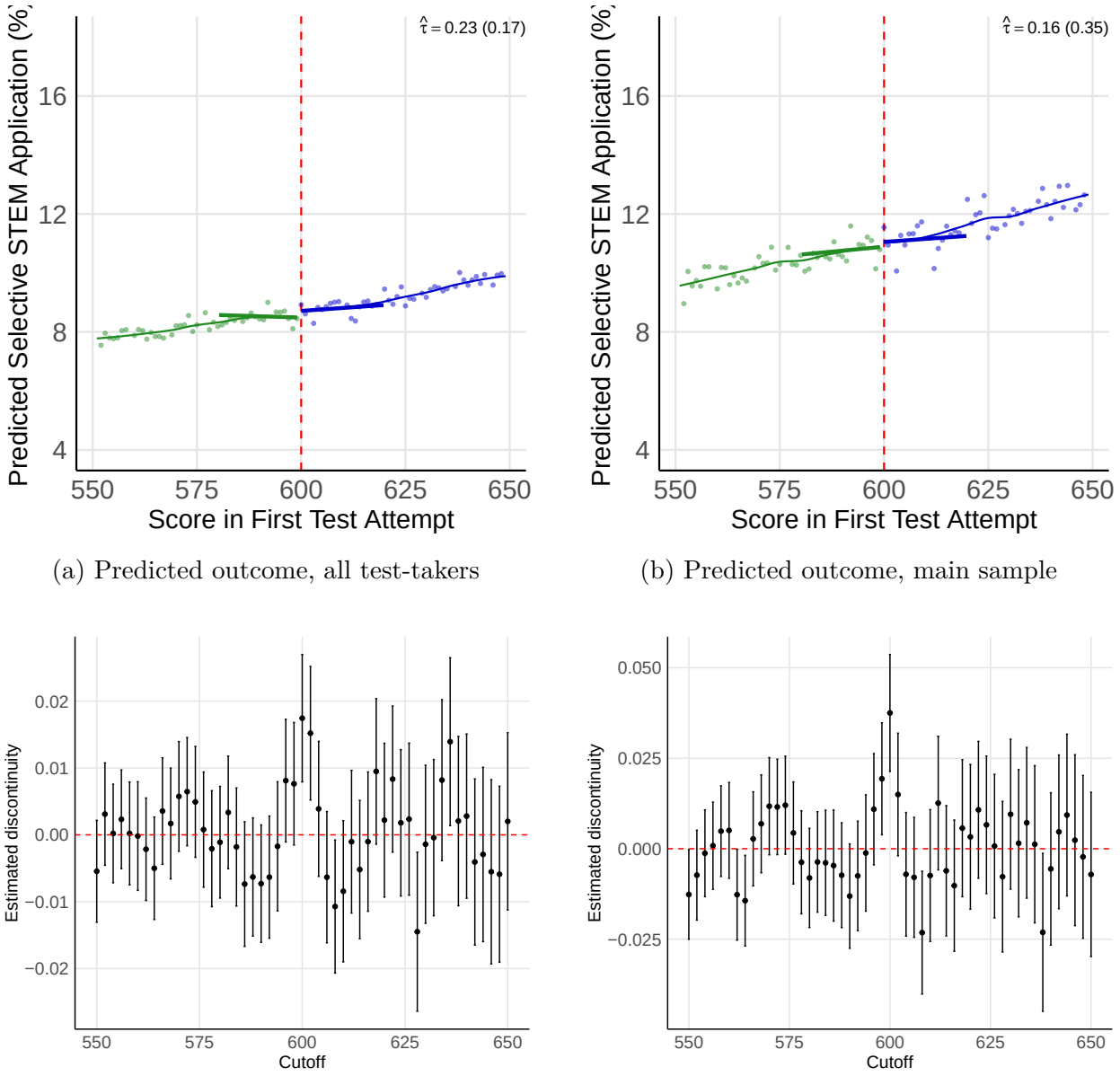
Figure A.2: Testing by Population Group



(a) Age at First Test (b) Number of Tests

Notes: The figures present the testing regularities of Jews and Arabs in Israel. The sample includes all individuals in Israel who participated in their first university entrance test during 2000–2009 (359,122 observations). Panel (a) shows the cumulative distribution of ages at the first test. Panel (b) shows the cumulative distribution of the number of tests taken.

Figure A.3: Regression Discontinuity Design Falsification Tests



(a) Predicted outcome, all test-takers

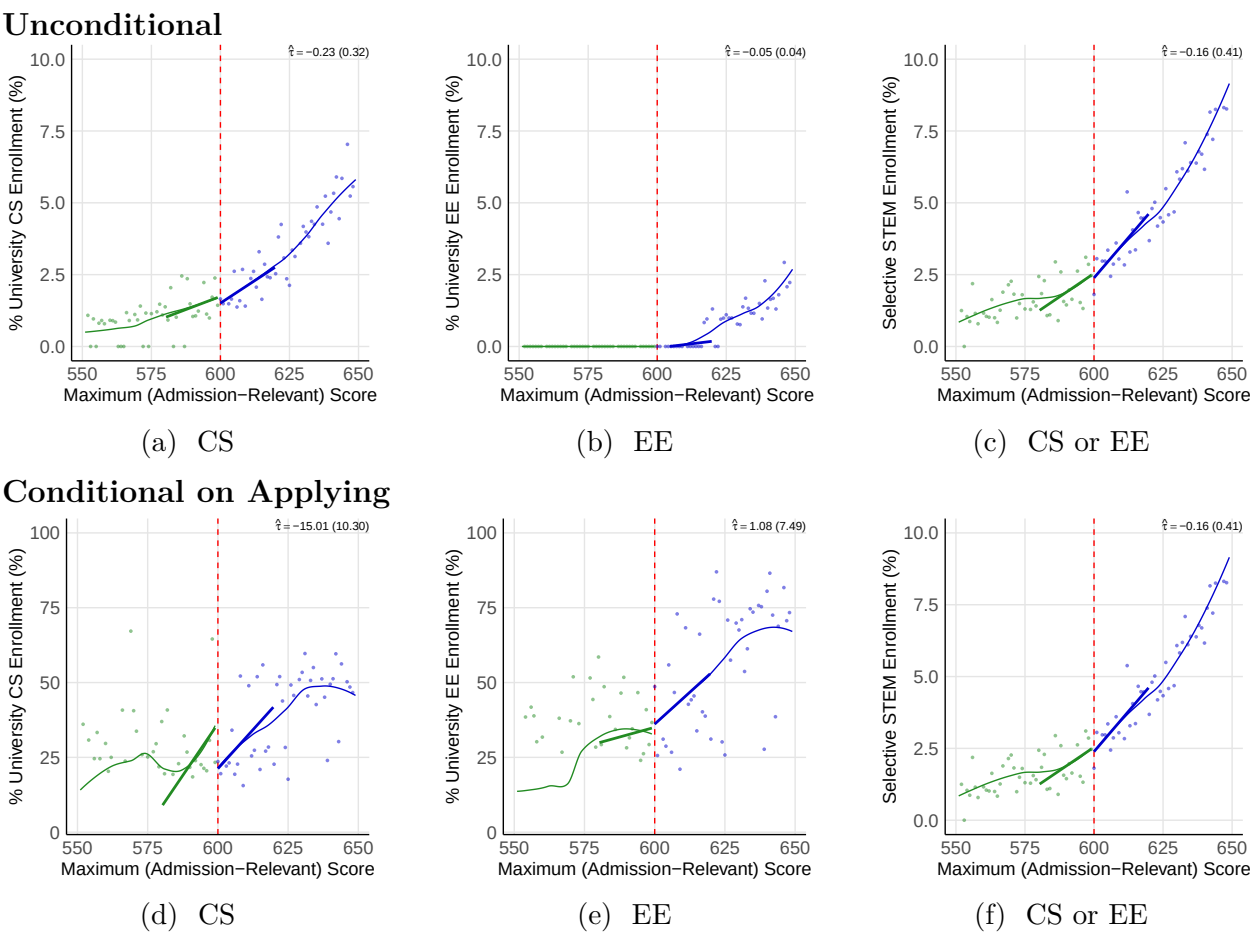
(b) Predicted outcome, main sample

(c) Non-round cutoffs, all test-takers

(d) Non-round cutoffs, main sample

Notes: These panels present falsification tests. Panels (a)–(b) show estimated discontinuities at the 600 cutoff in predicted application to selective STEM degrees, where predictions are based only on background characteristics. Panels (c)–(d) show estimated discontinuities at non-round score cutoffs, using selective STEM application as the outcome. Blue lines plot local-linear regression discontinuity fits from Equation 1; green lines show LOESS fits estimated over a broader set of scores. Estimates are based on first-time UPET takers in 2000–2009, restricted to those with first scores 580–619 (for panels a–b) or within 20 points of the tested cutoff (for panels c–d). Standard errors are clustered by score.

Figure A.4: Selective STEM Admissions

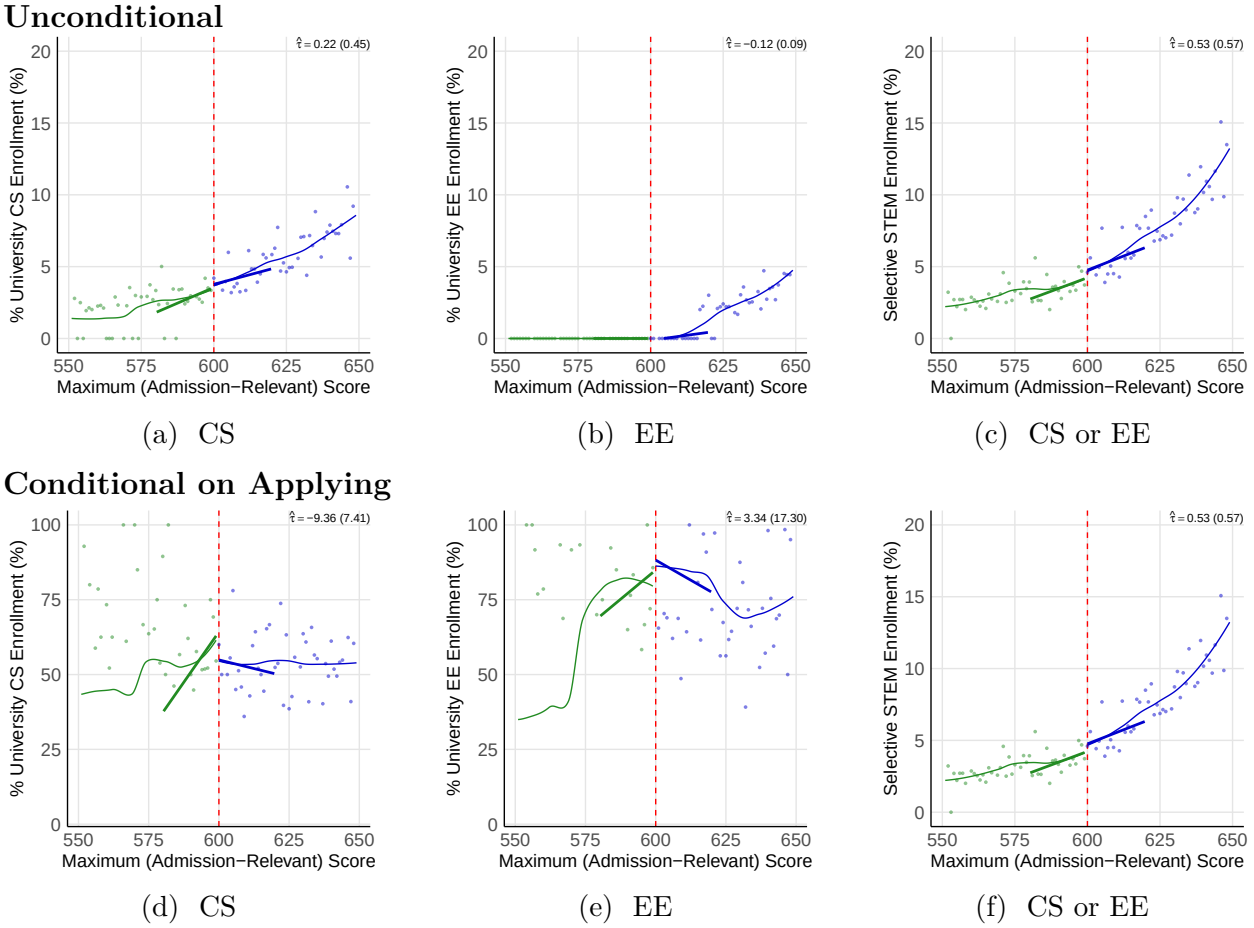


(a) CS (b) EE (c) CS or EE

(d) CS (e) EE (f) CS or EE

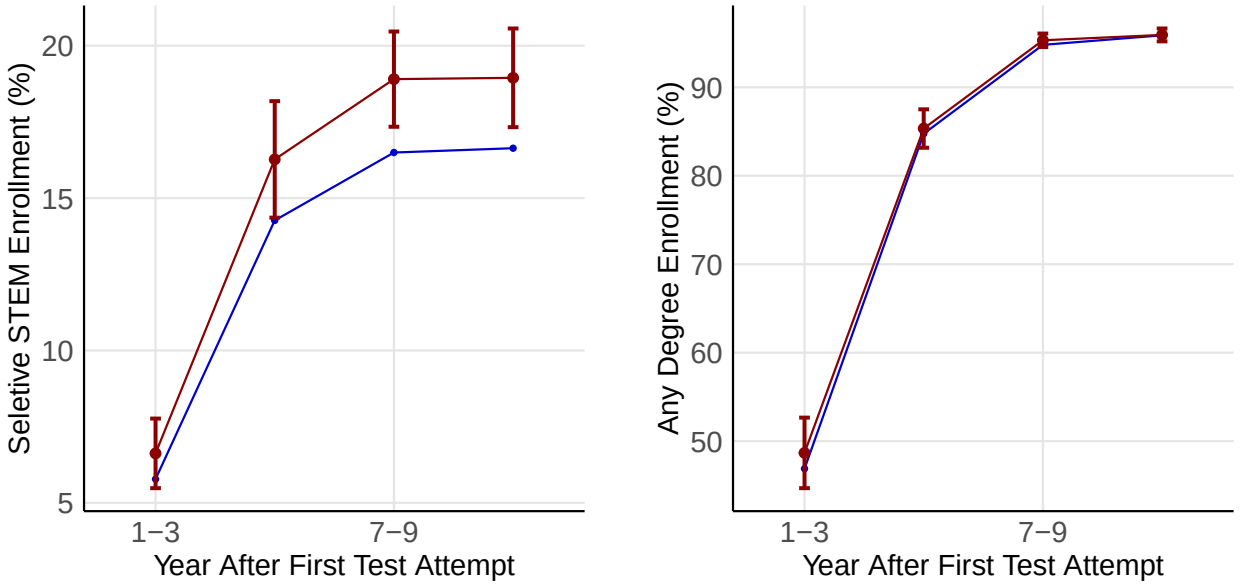
Notes: Panels (a) and (b) plot the relationship between attendance in selective STEM degrees (y-axis) and admission-relevant (highest-ever) university entrance test scores (x-axis). Panels (c) and (d) show the likelihood of attending these degrees, conditional on applying. Panels (a) and (b) include all individuals in Israel who took their first UPET between 2000 and 2009 ($N = 359,122$). Panels (c) and (d) are restricted to applicants to CS and EE degrees, respectively. "CS" denotes Computer Science; "EE" denotes Electrical Engineering.

Figure A.5: Selective STEM Admissions (Main Sample)



Notes: Panels (a) and (b) plot the relationship between attendance in selective STEM degrees (y-axis) and admission-relevant (highest-ever) university entrance test scores (x-axis). Panels (c) and (d) show the likelihood of attending these degrees, conditional on applying. Panels (a) and (b) include all individuals in Israel who took their first UPET between 2000 and 2009 ($N = 359,122$). Panels (c) and (d) are restricted to applicants to CS and EE degrees, respectively. "CS" denotes Computer Science; "EE" denotes Electrical Engineering.

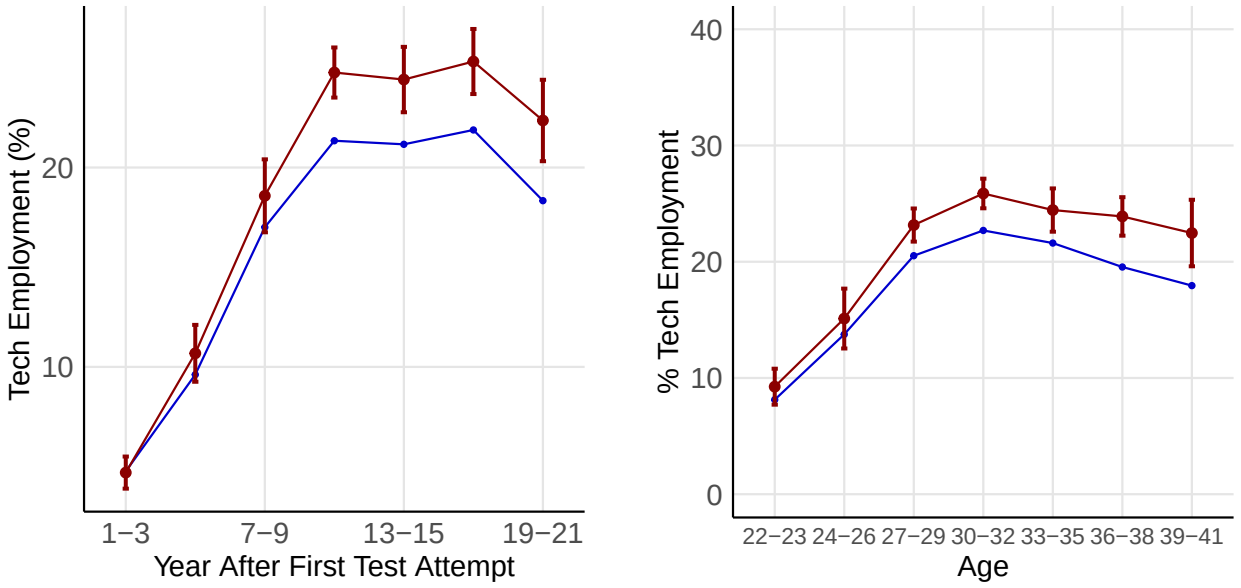
Figure A.6: Degree Enrollment Effects of Crossing 600 Over Time (Main Sample)



(a) STEM (%) (b) Any (%)

Notes: This figure shows the estimated effect of scoring just above 600 on degree enrollment over time. The x-axis indicates years since the first test; the y-axis values are computed from the estimation of Equation 1: blue dots plot the fitted level at 600^{-} (baseline), red dots plot baseline $+\tau$, i.e., the fitted level at 600^{+}. The red–blue gap equals τ at each threshold. Outcomes are enrollment in a selective STEM degree (panel a) and in any degree (panel b), measured 3, 6, 9, and 12 years after the test. Dots represent point estimates, and lines indicate 95 percent confidence intervals, based on robust standard errors clustered by score. The estimation sample includes 17,912 first-time test-takers (age 20 and below) in 2000–2009 who scored between 581 and 619.

Figure A.7: Effects of Crossing 600 on Tech Employment Over Time (Main Sample)

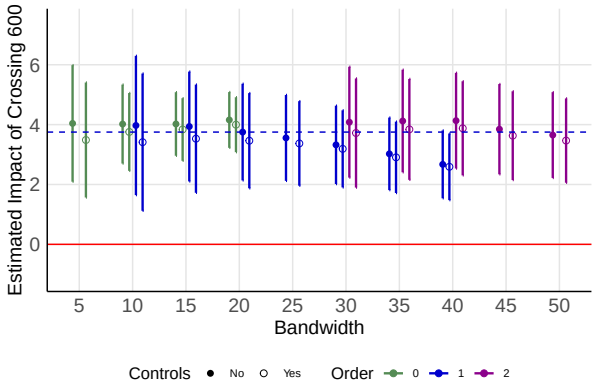


(a) Period relative to test

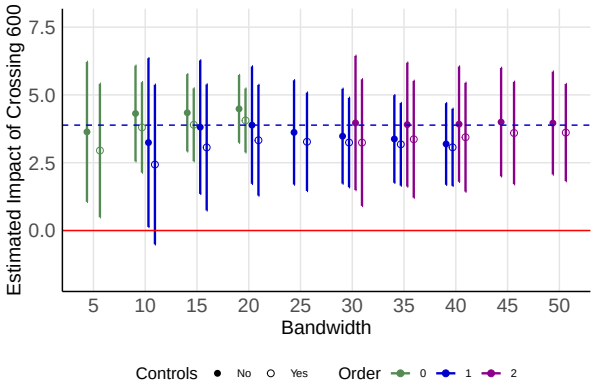
(b) Age

Notes: This figure shows baseline means (blue solid line) and implied means above the 600 cutoff (red solid line), where the latter equals the baseline mean plus the regression discontinuity estimate (using Equation 1). The outcomes are tech employment indicators measured in three-year bins after the first test (a) or age in three-year bins (b). Error bars display 95 percent confidence intervals for the regression discontinuity estimate, based on robust standard errors clustered by score. The estimation sample is the paper's main sample: 17,912 first-time test takers in 2000–2009, aged 20 or younger, restricted to scores 581–619 on the first test.

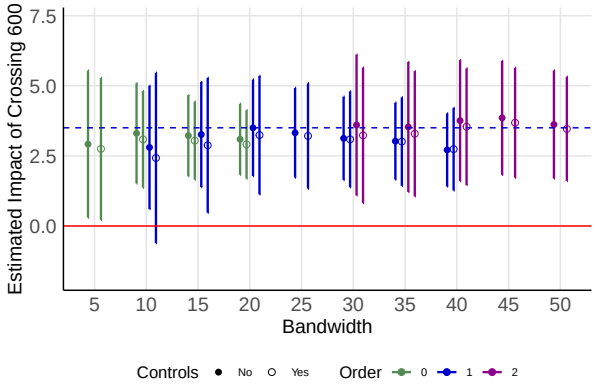
Figure A.8: Robustness of Regression Discontinuity Estimates (Main Sample)



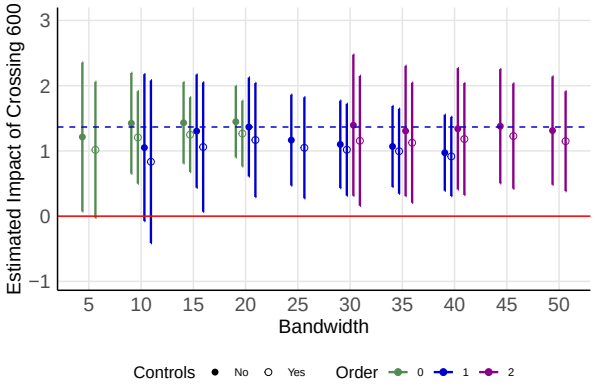
(a) Selective STEM Application (%)



(b) Lucrative STEM Enrollment (%)



(c) Tech Employment (%)

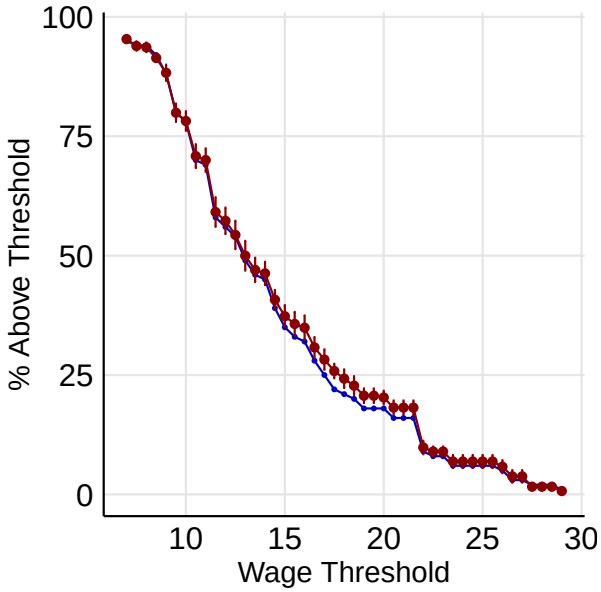


(d) Monthly Wage (NIS thousands)

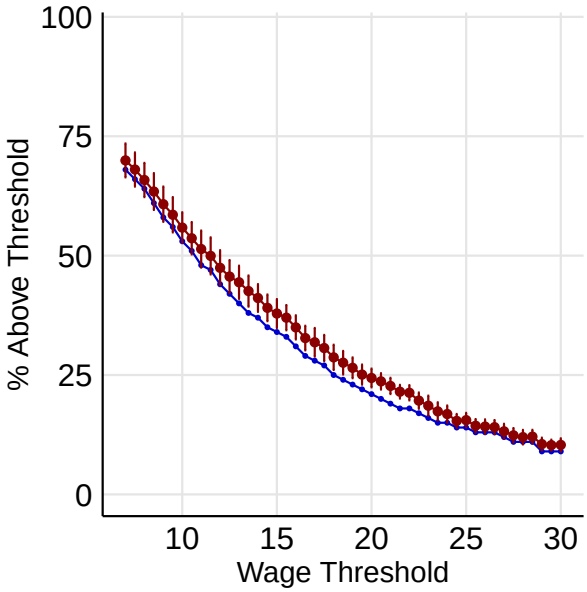
Notes: Each panel shows estimates of the impact of crossing the 600 score cutoff under alternative specifications. Points are regression discontinuity coefficients (τ from Equation 1) estimated using different bandwidths (x-axis), polynomial orders, and sets of control variables (detailed in the text). The sample includes 17,912 first-time test-takers in 2000–2009, age $\leq$20 at first attempt. Outcomes: (a) selective STEM application, (b) Lucrative STEM application, (c) tech employment, and (d) monthly wage. Employment and earnings are measured at 2022-2023.

Figure A.9: Effects of Crossing 600 on Degree-Linked and Individual Wage Thresholds

Main Sample

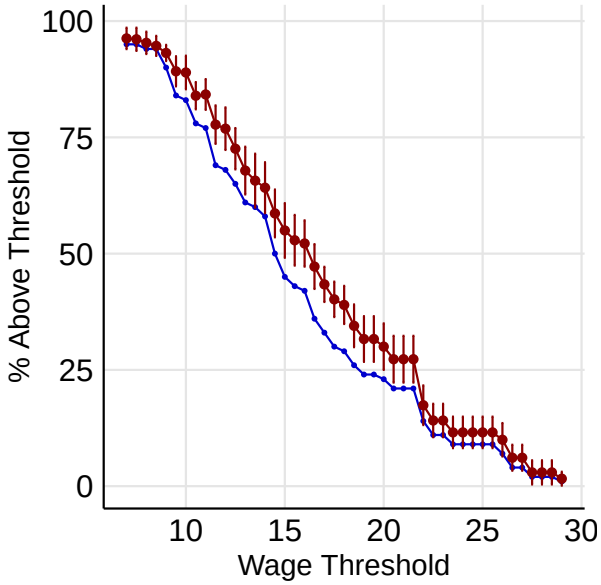


(a) Degree

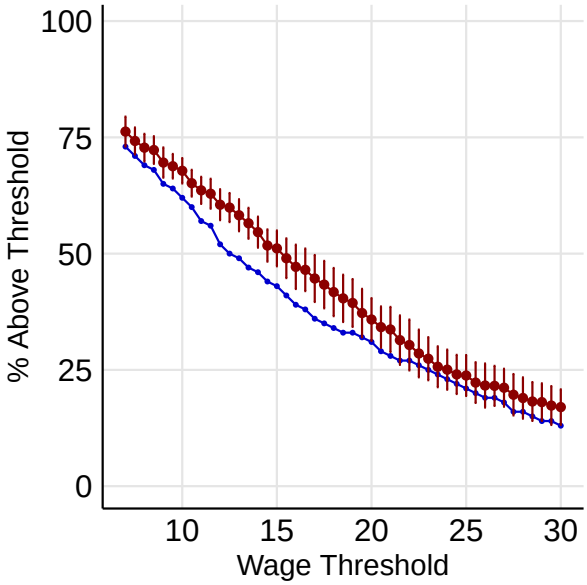


(b) Actual

Quantitatively Advantaged Sample



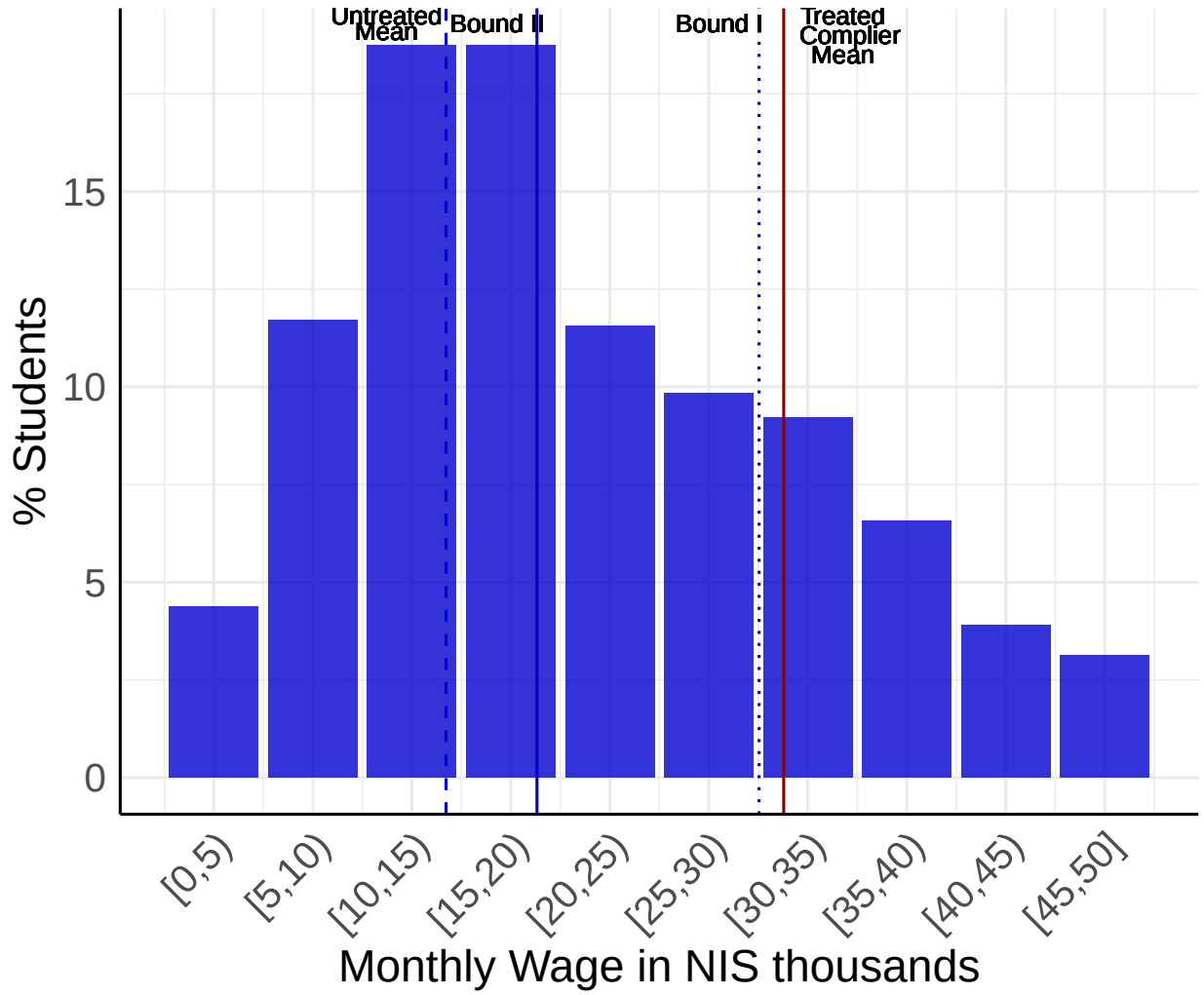
(c) Degree



(d) Actual

Notes: Each point is computed from the estimation of Equation 1: blue dots plot the fitted level at 600^- (baseline), red dots plot baseline $+\tau$, i.e., the fitted level at 600^+. The red–blue gap equals τ at each threshold. Panel (a): outcome is an indicator for selecting a degree with *associated* average earnings above the threshold on the x-axis. Panel (b): outcome is an indicator for the *individual's* own wage exceeding the threshold. Effects are measured in percentage points (y-axis). The sample includes first-time UPET takers in 2000–2009, age 20 or younger, with first-test scores between 580 and 619. Standard errors are robust and clustered by score.

Figure A.10: Constructing Bounds on the Returns to Degree Choice (Quantitatively Advantaged Sample)



Notes: The figure shows wage distributions for students not enrolled in any Lucrative STEM degree within the 591–599 window of first test-attempt scores (blue). As discussed in Appendix E, this sample consists of both untreated compliers and never-takers with respect to the instrumental-variables framework with a binary treatment. The vertical red line marks the estimated mean for treated compliers. I report three estimates for the counterfactual untreated-complier mean: (i) the overall mean in this group (untreated mean, solid blue line), (ii) the mean among those with the most rewarding degrees in this group (bound II, dashed blue line), and the mean among those with the highest wages in this group (bound I, dotted blue line). These estimates can be viewed as alternative conservative approaches to assessing local returns for compliers in the presence of binarization failures, as discussed in more detail in the appendix. The sample includes quantitatively advantaged test-takers who took the UPET for the first time between 2000 and 2009, at ages 20 and below, and who scored between 591 and 599. Wages are measured at ages 30–32 in thousands of NIS. Lucrative STEM degrees are defined in Table A.2.

Table A.1: Average Earnings by Postsecondary Degree and Industry

		Industry		Earnings (k NIS)			
	N	Industry	Share	Monthly	Annual	Father	Mother
A. No Postsecondary Degree							
Never enrolled	98082	Tech	8.1%	13.2	148.2	124.9	57.6
		Other	69.1%	9.2	105.1	109.5	40.5
Dropout	102786	Tech	18.5%	16.5	185.8	151.0	75.2
		Other	67.2%	8.7	98.2	128.7	55.9
B. Degree in Computer Science or Electrical Engineering							
College	18070	Tech	64.9%	23.6	268.6	167.8	81.7
		Other	25.8%	13.5	156.0	134.4	56.1
University	11814	Tech	79.4%	27.2	309.1	226.1	100.8
		Other	13.6%	14.7	171.9	202.1	94.8
Elite	15014	Tech	80.9%	30.2	342.3	253.5	117.6
		Other	8.9%	14.9	171.6	268.1	121.5
C. Degree in Other Fields							
College	221198	Tech	15.3%	15.5	172.5	179.5	84.9
		Other	78.5%	10.4	115.4	162.7	69.7
University	159552	Tech	19.2%	17.0	188.9	200.7	91.2
		Other	73.8%	10.6	117.5	164.0	73.3
Elite	116752	Tech	23.6%	19.5	217.7	234.8	108.3
		Other	66.8%	11.8	130.9	220.9	96.7

Note: This table reports average earnings by postsecondary degree and employment industry. Students are grouped by degree status, field, and institution type. N gives the number of students in each group. Within each group, students are split into those employed in the tech sector (2022–2023) and those employed in other industries; Share is the percentage employed in tech/other industries; the difference from 100 percent reflects the share not employed. The table then presents average monthly wages and annual earnings, along with average parental annual earnings (father and mother) when students were aged 14–16. All earnings are measured in thousands of NIS (2023 values). Selective STEM fields are computer science and electrical engineering. Tech includes various industries as detailed in the appendix.

Table A.2: Lucrative STEM Degrees

Field	Type	Students (1)	% Tech (2)	Wage (3)
Computer Engineering	Elite	965	71.6	29.5
Comp. Elec. Engineering	Elite	983	73.2	28.6
Computer Science	Elite	4,657	66.3	27.3
Electrical Engineering	Elite	4,833	66.1	26.3
Computer Engineering	Uni	1,068	74.0	25.6
Systems Engineering	Elite	420	69.0	23.4
Systems Engineering	Uni	671	72.7	23.4
Math & Comp Sci	Elite	232	55.8	23.3
Communications Engineering	Uni	567	75.2	23.1
Electrical Engineering	Uni	2,453	65.6	22.5
Computer Science	Uni	6,816	58.3	21.8
Computer Science	College	6,157	60.1	21.5
Industrial Engineering	Elite	2,060	54.6	20.4
Electronic Engineering	Uni	258	67.9	20.4
Physics	Elite	2,875	44.3	19.7
Math	Elite	1,488	40.7	18.9
Math	Uni	3,797	42.0	18.6
Aerospace Engineering	Elite	680	46.0	18.3
Industrial Engineering	Uni	2,572	50.1	18.0
Bio-Med Engineering	Elite	957	51.4	18.0
Computer Engineering	College	3,628	55.5	17.7

Notes: This table presents information on lucrative STEM degrees in Israel, defined as fields with more than 40% of graduates working in the tech sector and earning salaries above 17,500 NIS per month, both measured at ages 30–32. The table shows the number of graduates (column 1), the share of graduates working in the tech sector at ages 30–32 (column 2), and the average monthly wage at ages 30–32 in thousands of NIS (2023 values, column 3) for each field and institution type. The sample excludes students in the main analysis (first-test scores 580–620) to avoid mechanical overlap with the treatment sample.

Table A.3: Falsification Tests: "Effects" of Crossing 600 on Basline Characteristics

	(1)	(2)	(3)	(4)	(5)
A. Individual					
	Age	% Female	% Israeli	% Religious	Order
	-0.02	-2.00	-0.49	-0.18	-1.15
	(0.08)	(1.25)	(0.54)	(0.86)	(0.90)
Mean	20.30	56.33	84.84	81.78	151.89
N	40,383	40,383	40,383	40,383	40,383
B. Test					
			Score by domain		
	Year	Month	Quantitative	English	Qualitative
	-0.04	-0.05	-0.15	0.45*	-0.01
	(0.03)	(0.08)	(0.20)	(0.26)	(0.18)
Mean	2004.63	7.06	118.57	119.26	115.63
N	40,383	40,383	40,383	40,383	40,383
C. Family					
		Mother		Father	
	Siblings	Educ	Income	Educ	Income
	-0.01	0.07	-3.22*	-0.01	0.57
	(0.04)	(0.06)	(1.85)	(0.07)	(5.03)
Mean	2.42	14.11	92.87	13.74	195.81
N	40,383	40,383	40,383	40,383	40,383

Notes: This table reports falsification tests for the regression discontinuity design, showing estimated discontinuous changes in baseline characteristics at the 600 cutoff on the first university entrance test. Columns (1)–(5) present estimates of τ from Equation 1, with robust standard errors in parentheses, clustered at the score level. The estimation sample includes first-time test takers in 2000–2009, with initial scores between 581 and 619. Earnings are measured in thousands of NIS (2023 values). Education is measured in years of schooling. All indicator outcomes are multiplied by 100. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Table A.4: Falsification Tests: "Effects" of Crossing 600 on Basline Characteristics (Main Sample)

	(1)	(2)	(3)	(4)	(5)
A. Individual					
	Age	% Female	% Israeli	% Religious	Order
	-0.02	-2.04	-0.83	-0.82	0.06
	(0.03)	(1.90)	(0.99)	(1.13)	(1.06)
Mean	18.17	61.50	77.60	75.56	156.44
N	17,912	17,912	17,912	17,912	17,912
B. Test					
			Score by domain		
	Year	Month	Quantitative	English	Qualitative
	0.10	-0.13	0.23	0.25	-0.29
	(0.07)	(0.09)	(0.17)	(0.47)	(0.25)
Mean	2003.61	8.38	119.65	119.35	114.51
N	17,912	17,912	17,912	17,912	17,912
C. Family					
		Mother		Father	
	Siblings	Educ	Income	Educ	Income
	-0.02	0.02	-4.26*	-0.02	1.33
	(0.06)	(0.08)	(2.23)	(0.15)	(3.80)
Mean	2.34	14.62	99.81	14.08	204.19
N	17,912	17,912	17,912	17,912	17,912

Notes: This table reports falsification tests for the regression discontinuity design, showing estimated discontinuous changes in baseline characteristics at the 600 cutoff on the first university entrance test. Columns (1)–(5) present estimates of τ from Equation 1, with robust standard errors in parentheses, clustered at the score level. The estimation sample is the paper's main sample: first-time test takers in 2000–2009, aged 20 or below, restricted to scores 581–619 on the first test. Earnings are measured in thousands of NIS (2023 values). Education is measured in years of schooling. All indicator outcomes are multiplied by 100. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Table A.5: Predictors for Selective STEM Applications

Variable	Estimate	(Std. Error)
Intercept	-2.44***	(0.24)
Female	-0.09***	(0.03)
Female	-1.46***	(0.02)
Regular school	-0.12***	(0.02)
Born in Israel	0.11***	(0.02)
Parents Born in Israel	0.04***	(0.02)
First-born	0.14***	(0.02)
Second-born	0.07***	(0.02)
Age < 18	0.23***	(0.02)
Age < 20	0.48***	(0.02)
Age < 24	0.50***	(0.03)
1 Sibling	-0.14***	(0.04)
2 Siblings	-0.14***	(0.04)
>2 Siblings	-0.24***	(0.04)
Father's income 0	0.13***	(0.03)
Father's income $<$ 50K	-0.02	(0.03)
Father's income $<$ 100K	-0.02	(0.02)
Father's income $<$ 200K	-0.22***	(0.02)
Father's income $<$ 1M	0.04	(0.06)
Mother's income 0	0.01	(0.02)
Mother's income $<$ 50K	0.02	(0.02)
Mother's income $<$ 100K	-0.10***	(0.02)
Mother's income $<$ 200K	-0.14***	(0.03)
Mother's income $<$ 1M	-0.42*	(0.22)
Father 12y education	0.09***	(0.02)
Father 13-15y education	0.31***	(0.02)
Father >15y education	0.52***	(0.02)
Mother 12y education	0.16***	(0.03)
Mother 13-15y education	0.39***	(0.03)
Mother >15y education	0.50***	(0.03)

Notes: This table reports regression estimates from the model used to predict the main outcome—an indicator for applying to selective STEM within five years of the baseline test. The sample includes all individuals in the main estimation sample whose initial test score was just below 600. Students in the main analysis group (first-test scores 580–620) are excluded to avoid mechanical overlap with the treatment sample. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Table A.6: Effects of Crossing 600 on University Applications, Subsamples Outside Main Analysis

	Selective STEM	Lucrative STEM	Other Mid	Other Low	Any	Associated Wage
	(1)	(2)	(3)	(4)	(5)	(6)
	A. Older (21 and Above)					
	-0.59	-0.99	-0.85	-0.46	-2.30*	-0.08
	(0.43)	(0.71)	(1.40)	(0.99)	(1.28)	(0.19)
Mean	4.93	9.76	20.49	25.04	55.29	13.96
N	20,672	20,672	20,672	20,672	20,672	11,245
	B. Arabs					
	9.34**	-0.68	3.53	-2.39	0.46	0.96
	(4.71)	(4.89)	(4.14)	(3.39)	(2.75)	(0.65)
Mean	23.87	44.79	35.02	10.79	90.60	18.48
N	1,799	1,799	1,799	1,799	1,799	1,639

Notes: This table reports the estimated effect of crossing the 600 threshold on university applications within five years of the first test for subsamples outside the main analysis. Columns (1)–(6) report coefficients τ from Equation 1; robust standard errors are in parentheses and clustered by score. Outcomes include indicators for applications to: selective STEM (computer science and electrical engineering in universities); lucrative STEM (listed in Table A.2); non-STEM degrees classified by associated salaries; Any; and the expected wage at ages 30–32 associated with the chosen degree (thousand NIS, 2023). indicator outcomes are multiplied by 100. The baseline sample includes first-time UPET takers in 2000–2009 with first scores 581–619. Results are shown in seperate panels for different subsamples: Panel A shows results for older test-takers (age ≥ 21); Panel B shows results for Arab test-takers. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Table A.7: Effects of Crossing 600 on University Applications, Alternative Outcomes

	CS	EE	Tech	STEM Ever
	(1)	(2)	(3)	(4)
	A. Main			
	3.15***	0.82*	5.03***	3.36***
	(0.48)	(0.43)	(0.73)	(0.94)
Mean	6.20	5.20	15.06	17.64
N	17,912	17,912	17,912	17,912
	B. Main, Quantitative Advantage			
	5.66***	3.21***	11.21***	8.48***
	(1.48)	(0.95)	(2.54)	(3.01)
Mean	8.22	8.77	19.96	25.21
N	4,623	4,623	4,623	4,623
	C. Main, No Quantitative Advantage			
	2.36***	0.09	3.08***	1.78
	(0.55)	(0.41)	(0.70)	(1.20)
Mean	5.52	4.01	13.42	15.10
N	13,289	13,289	13,289	13,289

Notes: This table reports the estimated effect of crossing the 600 score cutoff on alternative university application outcomes. Columns (1)–(4) show the estimated coefficients τ from Equation 1, with robust standard errors in parentheses, clustered at the score level. Outcomes include indicators for applying to computer science, electrical engineering, any degree associated with high tech employment rate (above 40%) within five years, an indicator for ever applying to any selective STEM degree (computer science and electrical engineering in universities). Dummies are multiplied by 100. The baseline sample includes all first-time UPET takers in Israel from 2000 to 2009 who scored between 580 and 620 on their first test. Panels A–C report results for the main sample and subsamples defined by quantitative advantage. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Table A.8: Effects of Crossing 600 on Admission-Relevant Outcomes

	Maximum Score (1)	Retest (2)	STEM× Retest (3)	STEM× Max>640 (4)	Matriculation Credits (5)	Mat. STEM Credits (6)
	A. Main					
	0.14	-2.41	2.12***	1.72**	0.41***	0.11
	(1.07)	(1.88)	(0.56)	(0.67)	(0.11)	(0.08)
Mean	628.94	49.53	7.04	4.80	28.86	5.29
N	17,912	17,912	17,912	17,912	17,828	17,828
	B. Main, Quantitative Advantage					
	5.66***	4.36	4.99**	5.33***	0.69***	0.56***
	(1.24)	(4.05)	(2.05)	(1.64)	(0.26)	(0.21)
Mean	627.60	51.31	9.75	4.48	29.56	6.27
N	4,623	4,623	4,623	4,623	4,613	4,613
	C. Main, No Quantitative Advantage					
	-1.66	-4.58***	1.23***	0.54	0.32***	-0.03
	(1.20)	(1.54)	(0.46)	(0.57)	(0.12)	(0.07)
Mean	629.40	48.94	6.13	4.91	28.62	4.96
N	13,289	13,289	13,289	13,289	13,215	13,215

Notes: This table reports the estimated effect of crossing the 600 score cutoff on admission-related outcomes. Columns (1)–(6) display estimated coefficients τ from Equation 1, with robust standard errors in parentheses, clustered at the score level. Outcomes include the application-relevant score, which is each student's maximum score achieved in this window (max score), an indicator for retaking the test; an indicator for the interaction between selective STEM application and retaking the test; an indicator for the interaction between selective STEM application and achieving a maximum score above 640 through retesting; total matriculation (Bagrut) credits; and total matriculation credits in scientific subjects (math, physics, biology, chemistry, computer science). Dummies are multiplied by 100. The estimation sample in panel A is the main sample, including first-time UPET takers from 2000 to 2009 at age 20 and below who scored between 580 and 620 on their first test. Panels B–C report results for subsamples defined by quantitative advantage (having a quantitative score above scores in other domains by at least five points); selective STEM degrees include computer science and electrical engineering in universities. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Table A.9: Effects of Crossing 600 on Young Siblings UPET Testing

	3 Years	5 Years	7 Years	Total
	(1)	(2)	(3)	(4)
	A. Main			
	4.08**	2.98	2.31	0.07
	(1.93)	(2.35)	(2.48)	(0.06)
Mean	15.12	31.63	48.21	0.94
N	8,405	8,405	8,405	8,405
	B. Main, Quantitative Advantage			
	2.10	-1.30	0.48	0.06
	(3.53)	(4.55)	(6.94)	(0.11)
Mean	16.91	32.66	47.93	1.04
N	2,267	2,267	2,267	2,267
	C. Main, No Quantitative Advantage			
	4.77***	4.45**	2.93*	0.08*
	(1.84)	(1.93)	(1.66)	(0.05)
Mean	14.48	31.27	48.31	0.90
N	6,138	6,138	6,138	6,138
	D. Arabs			
	3.21	5.81	2.43	0.18
	(7.93)	(6.81)	(6.42)	(0.17)
Mean	45.61	62.06	68.18	2.01
N	1,269	1,269	1,269	1,269
	E. Older (21 and Above)			
	3.02	1.88	1.02	0.04
	(1.91)	(2.77)	(2.28)	(0.04)
Mean	22.97	40.02	46.07	0.82
N	9,543	9,543	9,543	9,543

Notes: This table reports the estimated effect of crossing the 600 score cutoff on young siblings' testing in the university entrance test. Columns (1)–(4) display estimated coefficients τ from Equation 1, with robust standard errors in parentheses, clustered at the score level. Outcomes include an indicators for any test taken by young siblings and the total number of tests taken by younger siblings. Dummies are multiplied by 100. The sample includes all first-time UPET takers in Israel from 2000 to 2009 who scored between 580 and 620 on their first test with any young sibling. ls A–C report results for the main sample and subsamples defined by quantitative advantage; Panels D–E present results for older and Arab test-takers. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Table A.10: Heterogeneity of the Effects of Crossing 600 on Degree Enrollment (Main Sample)

	Difference	Z=0	Z=1
	(1)	(2)	(3)
	Z: Quantitative Advantage		
Selective STEM	4.90**	0.65	5.55***
	(2.43)	(0.61)	(2.15)
		[9.57]	[13.94]
Lucrative STEM	8.24***	1.92*	10.17***
	(2.53)	(1.03)	(1.91)
		[19.42]	[29.67]
	Z: Female		
Selective STEM	-2.09	2.76*	0.67
	(1.87)	(1.49)	(0.65)
		[20.37]	[4.59]
Lucrative STEM	-4.35*	5.97***	1.62*
	(2.58)	(2.17)	(0.97)
		[37.44]	[12.30]
	Z: High-SES		
Selective STEM	1.57	0.92	2.48***
	(1.55)	(1.18)	(0.72)
		[9.37]	[11.70]
Lucrative STEM	0.14	3.75***	3.89**
	(2.41)	(1.29)	(1.57)
		[19.49]	[23.99]

Notes: This table reports the estimated heterogeneity of the effects of crossing the 600-score cutoff. Column (1) reports estimates of the difference in estimated discontinuity effects across subgroups defined by variable Z. Columns (2)–(3) report the estimated impacts (τ from Equation 1) for subgroups with different values of Z, separately. Estimaion is based on a local linear regression, fully interacted with the indicator for Z. Robust standard errors, clustered at the score level, are in parentheses. Outcomes are indicators for ever enrollment in a university selective STEM degree (computer science and electrical engineering) or in a lucrative STEM degree (listed in Table A.2). Heterogeneity dimensions include quantitative advantage (quantitative score at least five points above other domains), female indicator, and high-SES indicator (both parents with any postsecondary education). The estimation sample is the paper's main sample: first-time test takers in 2000–2009, aged 20 or younger, restricted to scores 581–619 on the first test. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Table A.11: Effects of Crossing 600 on Postsecondary Degree Outcomes (with Quantitative Advantge)

	(1)	(2)	(3)	(4)	(5)
A. Ever Enrollment					
	Selective STEM	Lucrative STEM	Other Mid	Other Low	Any
	5.86***	10.17***	-2.61	-6.95**	0.60
	(2.10)	(1.94)	(2.83)	(2.72)	(0.98)
Mean	13.71	29.67	35.67	30.71	96.05
N	4,623	4,623	4,623	4,623	4,623
B. Completion					
	Selective STEM	Lucrative STEM	Other Mid	Other Low	Any
	3.93**	9.31***	-3.61	-8.24***	-2.53
	(1.97)	(2.58)	(3.02)	(3.12)	(1.86)
Mean	9.76	24.82	34.06	30.79	89.67
N	4,623	4,623	4,623	4,623	4,623

Notes: This table reports the estimated impact of crossing the 600-score cutoff on degree enrollment and completion. Columns (1)–(5) present estimates of τ from Equation 1, with robust standard errors in parentheses, clustered at the score level. The estimation sample is a subsample of the paper's main sample: first-time test takers in 2000–2009 with quantitative advantage, aged 20 or below, restricted to scores 581–619 on the first test. Panel A reports ever-enrollment outcomes, and Panel B reports degree completion. Other Mid and Other Low classify remaining degrees by associated salaries, as described in the text. Any refers to degrees in any field and institution type (universities or colleges). All indicator and rank outcomes are multiplied by 100. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Table A.12: Effects of Crossing 600 on Degree Achievement

	Degree GPA		UPET	
	Above Median (1)	Average (2)	Above Median (3)	Average (4)
	A. Main			
	1.70	-0.11	-2.46	1.75
	(1.88)	(0.07)	(1.67)	(2.12)
Mean	52.62	84.23	29.49	626.00
N	12,733	12,733	12,733	12,725
	B. Main, Quantitative Advantage			
	0.10	-0.18	-6.69**	9.38***
	(3.71)	(0.16)	(2.96)	(3.25)
Mean	55.64	83.69	24.02	631.27
N	3,489	3,489	3,489	3,487
	C. Main, No Quantitative Advantage			
	2.27	-0.09	-1.04	-0.90
	(2.17)	(0.07)	(1.96)	(2.42)
Mean	51.56	84.42	31.42	624.19
N	9,244	9,244	9,244	9,238

Notes: This table reports the estimated effect of crossing the 600 score cutoff on degree achievement. Columns (1)–(4) show the estimated coefficients τ from Equation 1, with robust standard errors in parentheses, clustered at the score level. Outcomes an indicator for GPA rank above median among degree graduates, the average GPA among degree graduates, an indicator for UPET rank above median among degree graduates, the average UPET among degree graduates. All outcomes are defined at the field–institution–cohort level. Indicator outcomes are multiplied by 100, such that coefficients represent precentage points. The baseline sample includes all first-time UPET takers in Israel from 2000 to 2009 who scored between 580 and 620 on their first test. Panels A–B report results for the two subsamples outside the main analysis: older test-takers and Arab test-takers. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Table A.13: Effects of Crossing 600 on Degree Enrollment, Subsamples Outside Main Analysis

	Selective STEM (1)	Lucrative STEM (2)	Other Mid (3)	Other Low (4)	Any (5)
	A. Arabs				
	0.83	-1.06	-1.31	4.80	2.42
	(3.28)	(4.26)	(3.75)	(5.55)	(2.08)
Mean	13.69	23.95	37.44	28.38	89.77
N	1,799	1,799	1,799	1,799	1,799
	B. Older				
	-0.58*	-0.57	-0.38	1.15	0.20
	(0.34)	(0.57)	(1.26)	(1.29)	(0.55)
Mean	2.53	9.40	28.37	56.60	94.37
N	20,672	20,672	20,672	20,672	20,672

Notes: This table reports the estimated effect of crossing the 600 score cutoff on degree enrollment. Columns (1)–(5) show the estimated coefficients τ from Equation 1, with robust standard errors in parentheses, clustered at the score level. All outcomes are multiplied by 100, such that coefficients represent precentage points. The baseline sample includes all first-time UPET takers in Israel from 2000 to 2009 who scored between 580 and 620 on their first test. Panels A–B report results for the two subsamples outside the main analysis: older test-takers and Arab test-takers. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Table A.14: Effects of Crossing 600 on Degree Enrollment, Additional Outcomes (Main Sample)

	(1)	(2)	(3)	(4)	(5)
A. Undergraduate Degrees					
	University (Any field)	Elite (Any field)	Elite (Sel. STEM)	Associated Wage	Associated Wage (Field)
	-3.59***	-2.80**	0.78	0.33**	0.38**
	(1.32)	(1.41)	(0.64)	(0.14)	(0.15)
Mean	75.60	37.87	4.85	14.47	14.95
N	17,912	17,912	17,912	17,167	17,167
B. Undergraduate Degree Fields					
	Social	Other STEM	Business& Econ	Para-Med	Humanities
	-3.07	-0.76	0.04	-1.10	-1.00
	(2.16)	(1.12)	(1.04)	(1.06)	(0.78)
Mean	32.47	21.25	11.32	8.80	15.21
N	17,912	17,912	17,912	17,912	17,912
C. Graduate and Professional Degrees					
	MA	MA (Uni)	MA (Elite)	Medicine	Law
	-1.00	-0.35	0.47	0.29	-0.98
	(1.52)	(1.31)	(1.43)	(0.34)	(0.73)
Mean	43.26	35.57	17.10	1.81	8.27
N	17,912	17,912	17,912	17,912	17,912

Notes: This table reports the estimated impact of crossing the 600 score cutoff on degree enrollment by field. Columns (1)–(6) report regression discontinuity estimates of τ from Equation 1, with robust standard errors in parentheses, clustered at the score level. The estimation sample is the paper's main sample: first-time test takers in 2000–2009, aged 20 or below, restricted to scores 581–619 on the first test. Panel A-B outcomes are indicators for ever enrollment in undergraduate degrees by institution type and field, as well as the average wage associated with the enrolled degree, computed out of sample as described in the text; Panel C outcomes are indicators for ever enrollment in advanced degrees, as well as in professional degrees: Medicine and Law. Indicator variables are multiplied by 100. Sel. STEM includes computer science and electrical engineering; Elite universities are the Hebrew University of Jerusalem, Tel Aviv University, and the Technion. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Table A.15: Effects of Crossing 600 on Degree Enrollment, Additional Outcomes (with Quantitative Advantge)

	(1)	(2)	(3)	(4)	(5)
A. Undergraduate Degrees					
	University (Any field)	Elite (Any field)	Elite (Sel. STEM)	Associated Wage	Associated Wage (Field)
	1.56	2.48	2.62*	1.23***	1.03***
	(2.75)	(3.15)	(1.46)	(0.30)	(0.34)
Mean	75.08	36.10	7.02	15.76	16.46
N	4,623	4,623	4,623	4,462	4,462
B. Undergraduate Degree Fields					
	Social	Other STEM	Business& Econ	Para-Med	Humanities
	-4.11*	-2.81	1.72	-3.47***	2.16
	(2.47)	(2.65)	(1.71)	(1.32)	(1.81)
Mean	25.67	23.36	10.96	7.45	7.40
N	4,623	4,623	4,623	4,623	4,623
C. Graduate and Professional Degrees					
	MA	MA (Uni)	MA (Elite)	Medicine	Law
	-2.00	-0.39	2.57	0.41	1.02
	(2.30)	(2.05)	(1.69)	(0.53)	(0.94)
Mean	44.16	36.34	16.08	1.37	5.29
N	4,623	4,623	4,623	4,623	4,623

Notes: This table reports the estimated impact of crossing the 600 score cutoff on degree enrollment by field. Columns (1)–(3) report regression discontinuity estimates of τ from Equation 1, with robust standard errors in parentheses, clustered at the score level. The estimation sample is a subsample of the paper's main sample: first-time test takers in 2000–2009 with quantitative advantage, aged 20 or below, restricted to scores 581–619 on the first test. Panel A-B outcomes are indicators for ever enrollment in undergraduate degrees by institution type and field, as well as the average wage associated with the enrolled degree, computed out of sample as described in the text; Panel C outcomes are indicators for ever enrollment in advanced degrees. Indicator variables are multiplied by 100. Sel. STEM includes computer science and electrical engineering; Elite universities are the Hebrew University of Jerusalem, Tel Aviv University, and the Technion. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Table A.16: Heterogeneity of the Effects of Crossing 600 on Labor-Market Outcomes (Main Sample)

	Difference	Z=0	Z=1
	(1)	(2)	(3)
	Z: Quantitative Advantage		
Monthly Wage	3.22*	0.58	3.80***
	(1.72)	(0.72)	(1.18)
		[23.20]	[26.84]
Annual Earnings	33.25**	5.42	38.67***
	(16.11)	(6.15)	(13.37)
		[213.17]	[255.52]
	Z: Female		
Monthly Wage	-1.20	1.71***	0.51
	(0.81)	(0.59)	(0.41)
		[31.97]	[19.68]
Annual Earnings	-21.80***	24.08***	2.28
	(7.95)	(7.09)	(3.74)
		[288.30]	[183.44]
	Z: High-SES		
Monthly Wage	0.82	0.91	1.73***
	(0.93)	(0.65)	(0.56)
		[24.03]	[24.24]
Annual Earnings	-2.51	14.65***	12.14**
	(7.22)	(5.39)	(5.91)
		[223.38]	[224.13]

Notes: This table reports the estimated heterogeneity of the effects of crossing the 600 score threshold on the first university entrance test. Column (1) reports estimates of the difference in estimated discontinuity effects across subgroups defined by variable Z. Columns (2)–(3) report the estimated impacts (τ from Equation 1) for subgroups with different values of Z, separately. Estimaion is based on a local linear regression, fully interacted with the indicator for Z. Robust standard errors, clustered at the score level, are in parentheses. Outcomes are an monthly wage and total income in 2022-2023 (in thousands of NIS 2023). Heterogeneity dimensions include quantitative advantage (quantitative score at least five points above other domains), female indicator, and high-SES indicator (both parents with any postsecondary education). The estimation sample is the paper's main sample: first-time test takers in 2000–2009, aged 20 or younger, restricted to scores 581–619 on the first test. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Table A.17: Effects of Crossing 600 on Labor-Market Outcomes in 2022-2023 (with Quantitative Advantge)

	(1)	(2)	(3)	
A. Earnings (NIS Thousands)				
	Annual Earnings	Log Annual	Monthly Wage	Log Monthly
	38.67***	11.53**	3.80***	12.88**
	(13.59)	(5.69)	(1.20)	(5.21)
Mean	255.52	1238.56	26.84	998.72
N	4,623	3,856	3,664	3,664
B. Top Earners				
	Rank	Top 10%	Top 5%	Top 1%
	4.03***	5.40**	4.25**	1.83***
	(1.13)	(2.42)	(1.79)	(0.45)
Mean	57.73	18.54	9.89	0.14
N	4,014	4,623	4,623	4,623
C. Employment				
	Any	Salaried	Self	Tech
	1.73	0.53	-0.24	7.58**
	(2.01)	(2.59)	(1.30)	(2.97)
Mean	85.93	79.20	10.00	31.12
N	4,623	4,623	4,623	4,623

Notes: This table reports the estimated impact of crossing the 600 score cutoff on labor-market outcomes in 2023. Columns (1)–(3) report the regression discontinuity estimates of τ from Equation 1, with robust standard errors in parentheses, clustered at the score level. The estimation sample is a subsample of the paper's main sample: first-time test takers in 2000–2009 with quantitative advantage, aged 20 or below, restricted to scores 581–619 on the first test. Earnings are measured in NIS (2023 values) thousands and other outcomes are measured as a percentage (multiplied by 100). $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Table A.18: Effects of Crossing 600 on Labor-Market Outcomes, Subsamples Outside Main Analysis

	Salaried Employment (1)	Self Employment (2)	Tech Employment (3)	Monthly Wage (4)	Annual Earnings (5)
	A. Arabs				
	-1.48	2.87	0.39	1.00	17.00
	(2.98)	(2.33)	(3.10)	(1.32)	(13.14)
Mean	81.78	12.72	13.21	20.04	190.31
N	1,799	1,799	1,799	1,471	1,799
	B. Older				
	-3.87***	-0.69	0.06	-0.29	-10.15***
	(1.13)	(0.90)	(1.15)	(0.34)	(3.86)
Mean	78.56	15.16	19.87	23.45	230.90
N	20,672	20,672	20,672	16,135	20,672

Notes: This table reports the estimated effect of crossing the 600 score cutoff on labor-market outcomes in 2022-2023. Columns (1)–(6) show the estimated coefficients τ from Equation 1, with robust standard errors in parentheses, clustered at the score level. The estimated coefficients are presented in either NIS thousands (2023 values) or percentage points (multiplied by 100). The baseline sample includes all first-time UPET takers in Israel from 2000 to 2009 who scored between 580 and 620 on their first test. Panels A–B report results for the two subsamples outside the main analysis: older test-takers and Arab test-takers. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Table A.19: Robustness of the Main Estimation Results

	Selective STEM Applications (%) (1)	Lucrative STEM Enrollment (%) (2)	Monthly Wage (NIS thousands) (3)	Tech Employment (%) (4)
A. MSE-optimal				
	3.99***	3.88***	1.36***	3.57***
	(0.51)	(0.77)	(0.37)	(0.86)
	[14]	[19]	[21]	[21]
N	13769	17912	15638	20036
B. Honest				
	3.96***	3.79**	1.55*	3.77**
	(1.19)	(1.52)	(0.73)	(1.45)
	[14]	[16]	[14]	[19]
N	13769	15284	10253	17912

This table reports the estimated effects of crossing the 600-score cutoff using alternative regression discontinuity methods. Panel A presents robust bias-corrected estimates following Calonico et al. (2015), and Panel B presents honest inference following Kolesár and Rothe (2018). The baseline sample is the paper's main sample: first-time test takers in 2000–2009, aged 20 or younger, restricted to scores 500–699 on the first test. The estimation sample is further restricted to observations within the chosen bandwidth (reported in the table). Standard errors are shown in parentheses, the chosen bandwidths in brackets, and the effective sample size is shown below. Outcomes include indicators for selective STEM application (computer science and electrical engineering in universities), lucrative STEM enrollment (fields listed in Table A.2), tech employment, and monthly wage (in thousands of 2023 NIS). Standard errors are shown in parentheses, the chosen bandwidths in brackets, and the effective sample size shown below. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Table A.20: Robustness of the Heterogeneity Estimation, by Relative Quantitative Advantage

	(1)	(2)	(3)
	Low	Medium	High
Selective STEM Applications (%)			
	2.74	1.72	10.77***
	(1.84)	(1.83)	(3.63)
N	5,448	11,958	4,623
Bandwidth	22.6	26.7	19.6
Lucrative STEM Enrollment (%)			
	0.83	2.08	13.67***
	(2.43)	(2.17)	(4.17)
N	6,803	15,461	5,431
Bandwidth	28.3	34.8	23.0
Tech Employment (%)			
	-6.00*	4.52**	9.11**
	(3.18)	(2.22)	(4.18)
N	6,103	14,969	5,190
Bandwidth	25.5	34.0	21.7
Monthly Wage (NIS thousands)			
	-758	748	4,601***
	(1,213)	(912)	(1,617)
N	5,460	13,383	5,042
Bandwidth	29.5	38.5	26.4

Notes: This table reports the robust bias-corrected estimated heterogeneous effects of crossing the 600 score threshold, using the algorithm developed by Calonico et al. (2025). The baseline sample is the paper's main sample: first-time test takers in 2000–2009, aged 20 or younger, restricted to scores 500–699 on the first test. The estimation sample is further restricted to observations within the chosen bandwidth (reported in the table). "Low" indicates students whose quantitative score is lower than both other test domains; "High" indicates students whose quantitative score exceeds both other domains by at least five points; "Mid" indicates all others. Outcomes include indicators for selective STEM application (computer science and electrical engineering in universities), lucrative STEM enrollment (fields listed in Table A.2), tech employment, and monthly wage (in thousands of 2023 NIS). $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Table A.21: Returns to Lucrative STEM Degrees, Local Randomization Approach

	Never	Compliers		Always	
	E[Y(0)]	E[Y(0)]	E[Y(1)]	E[Y(1)]	**LATE**
	(1)	(2)	(3)	(4)	**(5)**
A. Sample: Main					
2022-2023 (NIS Thousands)	21.0	15.2	41.5	36.7	**26.29*** **
N=7,064, F=27.7	(0.2)	(4.4)	(4.1)	(0.7)	**(6.07)**
2022-2023 (Log)	9.7	9.6	10.5	10.3	**0.81*** **
N=7,064, F=27.7	(0.0)	(0.2)	(0.2)	(0.0)	**(0.22)**
30-32 (NIS Thousands)	13.0	9.5	28.7	23.2	**19.21*** **
N=7,734, F=22.8	(0.2)	(3.7)	(2.2)	(0.4)	**(3.09)**
30-32 (Log)	9.3	9.1	10.1	9.8	**0.97*** **
N=7,734, F=22.8	(0.0)	(0.3)	(0.1)	(0.0)	**(0.26)**
33-35 (NIS Thousands)	16.5	14.7	31.9	29.6	**17.22*** **
N=6,968, F=26.0	(0.2)	(3.8)	(2.9)	(0.4)	**(4.33)**
33-35 (Log)	9.5	9.4	10.1	10.1	**0.72** **
N=6,968, F=26.0	(0.0)	(0.2)	(0.1)	(0.0)	**(0.27)**
B. Sample: Quantitatively-Advantaged					
2022-2023 (NIS Thousands)	24.0	14.7	45.4	37.0	**30.73*** **
N=1,863, F=21.8	(0.7)	(4.9)	(4.3)	(0.8)	**(7.49)**
2022-2023 (Log)	9.9	9.6	10.6	10.4	**1.00** **
N=1,863, F=21.8	(0.0)	(0.2)	(0.2)	(0.0)	**(0.36)**
30-32 (NIS Thousands)	14.7	11.0	31.3	24.9	**20.30*** **
N=2,000, F=17.1	(0.2)	(3.3)	(4.9)	(0.9)	**(5.14)**
30-32 (Log)	9.4	9.2	10.1	9.9	**0.83*** **
N=2,000, F=17.1	(0.0)	(0.2)	(0.2)	(0.0)	**(0.24)**
33-35 (NIS Thousands)	19.4	11.5	34.5	31.0	**22.93*** **
N=1,770, F=20.5	(0.5)	(4.6)	(3.4)	(0.9)	**(4.86)**
33-35 (Log)	9.7	9.2	10.2	10.2	**0.98*** **
N=1,770, F=20.5	(0.0)	(0.2)	(0.2)	(0.0)	**(0.22)**

Notes: This table reports estimated returns to enrolling in lucrative STEM degrees (δ from Equation 2, estimated as a kernel-weighted difference in means at the cutoff). It also reports estimated mean outcomes by type: never-takers, compliers (treated and untreated), and always-takers. Outcomes are monthly wages in thousands of NIS (2023 values), measured in 2022–2023 and at ages 30–32 and 33–35. The endogenous variable is an indicator for ever enrolling in a lucrative STEM degree (fields listed in Appendix Table A.2). The estimation sample in Panel A is the main sample: first-time test takers in 2000–2009, aged 20 or below, with scores between 591 and 609 on the first test. Panel B further restricts the sample to quantitatively advantaged students, in order to reduce concerns about the validity of the underlying assumptions (see discussion in Section 5). $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Table A.22: Returns to Lucrative STEM Degrees in Annual Earnings

	Never	Compliers		Always	
	E[Y(0)]	E[Y(0)]	E[Y(1)]	E[Y(1)]	**LATE**
	(1)	(2)	(3)	(4)	**(5)**
A. Sample: Main					
2022-2023 (NIS Thousands)	196.1	156.4	496.9	341.9	**340.48*** **
N=17,912, F=12.4	(2.3)	(92.9)	(52.1)	(3.8)	**(110.81)**
2022-2023 (Log)	12.1	12.0	13.2	12.7	**1.19** **
N=14,797, F=12.6	(0.0)	(0.3)	(0.2)	(0.0)	**(0.50)**
30-32 (NIS Thousands)	123.2	65.7	344.4	236.3	**278.69*** **
N=17,912, F=12.4	(1.3)	(73.8)	(40.0)	(3.6)	**(68.52)**
30-32 (Log)	11.7	10.9	12.8	12.3	**1.93*** **
N=15,416, F=12.6	(0.0)	(0.4)	(0.2)	(0.0)	**(0.50)**
33-35 (NIS Thousands)	155.5	212.7	388.4	289.4	**175.73** **
N=16,937, F=11.7	(1.5)	(41.1)	(50.6)	(4.0)	**(66.03)**
33-35 (Log)	11.9	12.0	12.7	12.5	**0.73***
N=14,286, F=13.1	(0.0)	(0.4)	(0.3)	(0.0)	**(0.37)**
B. Sample: Quantitatively-Advantaged					
2022-2023 (NIS Thousands)	222.1	129.9	510.3	353.7	**380.40*** **
N= 4,623, F=16.8	(7.3)	(94.6)	(85.0)	(7.0)	**(123.19)**
2022-2023 (Log)	12.3	12.1	13.1	12.7	**1.07** **
N= 3,856, F=15.7	(0.0)	(0.3)	(0.3)	(0.0)	**(0.52)**
30-32 (NIS Thousands)	142.9	64.1	341.4	251.3	**277.32*** **
N= 4,623, F=16.8	(2.8)	(53.1)	(50.9)	(7.1)	**(56.01)**
30-32 (Log)	11.8	11.4	12.4	12.3	**1.00*** **
N= 4,014, F=14.4	(0.0)	(0.3)	(0.2)	(0.0)	**(0.32)**
33-35 (NIS Thousands)	181.6	129.6	389.1	303.0	**259.51*** **
N= 4,280, F=15.1	(4.9)	(76.2)	(66.7)	(7.7)	**(70.56)**
33-35 (Log)	12.1	11.6	12.7	12.5	**1.03** **
N= 3,639, F=15.4	(0.0)	(0.4)	(0.2)	(0.0)	**(0.43)**

Notes: This table reports estimated returns to enrolling in lucrative STEM degrees (δ from Equation 2, estimated using local linear regression). It also reports estimated mean outcomes by type: never-takers, compliers (treated and untreated), and always-takers. Outcomes are total annual earnings in thousands of NIS (2023 values), measured in 2022–2023 and at ages 30–32 and 33–35. The endogenous variable is an indicator for ever enrolling in a lucrative STEM degree (fields listed in Appendix Table A.2). The estimation sample in Panel A is the main sample: first-time test takers in 2000–2009, aged 20 or below, with scores between 581 and 619 on the first test. Panel B further restricts the sample to quantitatively advantaged students, in order to reduce concerns about the validity of the underlying assumptions (see Section 5). $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Table A.23: Causal Effect of Lucrative STEM Degree Enrollment on Tech Employment

	Never	Compliers		Always	
	E[Y(0)]	E[Y(0)]	E[Y(1)]	E[Y(1)]	**LATE**
	(1)	(2)	(3)	(4)	**(5)**
A. Sample: Main					
30-32	14.1	9.8	92.7	54.2	**82.91*** **
N=17,912, F=12.1	(0.3)	(15.2)	(11.7)	(1.0)	**(15.37)**
33-35	14.4	6.4	81.5	52.4	**75.14*** **
N=16,937, F=11.3	(0.4)	(12.8)	(10.4)	(0.9)	**(14.88)**
2022-2023	18.1	4.1	95.4	57.1	**91.35*** **
N=17,912, F=12.1	(0.4)	(14.9)	(16.5)	(1.0)	**(19.69)**
B. Sample: Quantitatively-Advantaged					
30-32	16.6	11.6	88.9	53.9	**77.29*** **
N= 4,623, F=16.8	(0.9)	(16.0)	(16.3)	(2.0)	**(17.42)**
33-35	16.6	9.9	71.3	52.3	**61.42*** **
N= 4,280, F=15.1	(0.8)	(14.8)	(9.5)	(1.3)	**(17.47)**
2022-2023	20.8	17.7	92.3	57.1	**74.59** **
N= 4,623, F=16.8	(1.5)	(21.8)	(12.1)	(1.6)	**(28.21)**

Notes: This table reports estimated effects of enrolling in lucrative STEM degrees (δ from Equation 2, estimated using local linear regression). It also reports estimated mean outcomes by type: never-takers, compliers (treated and untreated), and always-takers. Outcomes are indicators for being employed in the tech industry, measured in 2022–2023 and at ages 30–32 and 33–35 (multiplied by 100). The endogenous variable is an indicator for ever enrolling in a lucrative STEM degree (fields listed in Appendix Table A.2). The estimation sample in Panel A is the main sample: first-time test takers in 2000–2009, aged 20 or below, with scores between 581 and 619 on the first test. Panel B further restricts the sample to quantitatively advantaged students, in order to reduce concerns about the validity of the underlying assumptions (see discussion in Section 5). $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Table A.24: Effects of Crossing 700 (Main Sample)

	(1)	(2)	(3)	(4)
A. University Applications and Admission Outcomes				
	Any	Lucrative STEM	Retesting	Max Score
	-0.10	1.01	0.01	0.64
	(1.04)	(1.78)	(1.73)	(0.77)
Mean	72.96	39.33	14.45	703.96
N	9,503	9,503	9,503	9,503
B. Degree Enrollment				
	Any	Lucrative STEM	Elite	Advanced
	1.35*	-0.11	0.45	8.06***
	(0.77)	(1.36)	(2.35)	(2.36)
Mean	97.19	43.62	66.68	49.21
N	9,503	9,503	9,503	9,503
C. Employment				
	Any	Self	Tech	Wage
	2.51**	-2.18	0.33	0.32
	(0.98)	(1.43)	(1.84)	(0.66)
Mean	72.67	12.72	37.41	33.26
N	9,503	9,503	9,503	7,194

Notes: This table reports the estimated impact of crossing the 700 Score cutoff. Columns (1)–(3) report the regression discontinuity estimates of τ from Equation 1, with robust standard errors in parentheses, clustered at the score level. The estimation sample is similar to the paper's main sample but focuses on a higher threshold: first-time test takers in 2000–2009, aged 20 or below, restricted to scores 680–720 on the first test. Earnings are measured in NIS (2023 values) thousands, scores range between 200 and 800, and other outcomes are measured as a percentage (multiplied by 100). $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Table A.25: Returns to Degree Choice Relative to Observational Differences

	Estimate
A. Sample: Main	
Level (Raw)	2.61***
N=14,932, F=5.3	(0.67)
Log (Raw)	1.08**
N=14,932, F=3.0	(0.50)
Log (Controls)	4.77**
N=14,932, F=2.3	(2.20)
Level (Controls)	4.19***
N=14,932, F=4.5	(1.09)
B. Sample: Quantitatively-Advantaged	
Level (Raw)	1.64***
N= 3,900, F=17.5	(0.47)
Log (Raw)	0.68**
N= 3,900, F=10.8	(0.30)
Log (Controls)	1.91***
N= 3,900, F=12.4	(0.59)
Level (Controls)	2.72***
N= 3,900, F=13.0	(0.69)

Notes: This table reports estimated returns relative to observational wage differences (δ^m from Equation A1, estimated by local linear regression). Two versions of the program premium are used: m_d, the unconditional mean wage of students in program d (denoted Raw), and m'_d, the program component from a saturated wage regression that adjusts for predetermined covariates (denoted Controls). Outcomes are monthly wages in thousands of NIS (2023 values), measured at ages 30–32 (the same period in which m is measured). The estimation sample in Panel A is the main sample: first-time test takers in 2000–2009, age 20 or below at the time of the test, with scores between 581 and 619. Panel B restricts the sample to quantitatively advantaged students to reduce concerns about the validity of the underlying assumptions (see Section 5).
starlegend

Appendix B Data Appendix

B.1 Data Sources

1) The National Institution for Testing and Evaluation provides information on the University Psychometric Entrance Test (UPET); It includes the scores and timing of all tests ever taken by each individual in the sample since 1995. 2) Higher Council of Education records of post-secondary degrees (enrollment and completions), the institution of study (colleges and universities), the field of study (one or two), and year; essential additional features of these data are the university application information. I observe these data for universities in all years and colleges in 2009 and later only. I also observe partial data on the admission decisions made for each application. 3) Israel Tax Authority (ITA) provides data on the earnings of employees and self-employed individuals from 2000-2023 (excluding 2021) and a three-digit code of industry of employment. 4) The population registry data includes a fictitious individual national ID number that appears in all the data sets described below and enables the matching and merging of the files at the personal level. It also contains information on the following student's family background variables: birth year, sex, locality, number of siblings, country of birth, and parental countries of birth. 5) The Ministry of Education has provided administrative data on Israeli high schools' universe since 1995. It provides data on students' matriculation programs and test scores (test scores are shown in bins).

B.2 Variable Definitions

University applications. The university application dataset is used to construct indicators for application behavior. These indicators capture applications submitted to different academic fields within five years after the test. Academic fields are classified using four-digit codes from the Central Bureau of Statistics classification; for example, computer science and electrical engineering correspond to codes 900 and 1020, respectively. Following Central Bureau of Statistics definitions, I group disciplines into broader categories, with the exception of lucrative STEM degrees, which are defined in the main text.

The dataset also contains partial information on admission decisions, with binary indicators for admission to a first-choice field and to any other field. Because these data are incomplete, the analysis focuses primarily on applications and enrollment outcomes rather than admissions.

Degree enrollment. I define an individual as enrolled in a degree if they are coded as enrolled in the enrollment dataset or if they completed the program in the degree attainment dataset. This combined definition provides a more comprehensive measure of enrollment, since the enrollment dataset contains some missing information on students' fields of study.

Degree GPA. The CBS provides near-complete data on degree GPA for graduates in my sample. For each individual who earned a degree and appears in this dataset, I assign their reported GPA. I then transform this measure into a percentile rank within groups defined by degree field (main field by institution type) and graduation year.

Tech employment. Employment in the tech sector is defined as working in either service or manufacturing firms classified as technology industries. Following the Central Bureau of Statistics definition, these include: pharmaceutical products for human and veterinary use; office and accounting machinery and computers; electronic components; electronic communication equipment; industrial control and supervision equipment; medical and scientific instruments; aircraft manufacturing; telecommunications; computer and related services; and research and development services.

Earnings outliers. To address earnings outliers, I restrict the sample in the primary analysis. All income measures are top-coded at the 99th percentile and limited to observations above half the minimum wage.

Young siblings' testing. I identify siblings by linking individuals who share the same mother in the population registry. For each test-taker in my sample, I retain one sibling only if the birth gap is no greater than five years; otherwise, the individual is excluded. For the retained cases, I record whether the sibling took the test within 3, 5, and 7 years, as well as the total number of tests ever taken.

Predicted likelihood of applying to selective STEM fields. I construct $\widehat{p}_i = \Pr(\text{apply to selective STEM within five years} \mid X_i)$ using a logistic regression estimated outside the regression discontinuity window. Specifically, I fit a logistic model of the outcome on pre-determined covariates, excluding observations with $s_i \in [580, 619]$. The covariates are indicators for Arab; female; non-religious (regular) school; born in Israel; parents born in Israel; first-born and second-born (family_order $= 1, 2$); age indicators $\mathbf{1}\{\text{age} < 18\}$, $\mathbf{1}\{\text{age} < 20\}$, $\mathbf{1}\{\text{age} < 24\}$; sibling counts ($\mathbf{1}\{\text{siblings} = 1\}$, $\mathbf{1}\{\text{siblings} = 2\}$, $\mathbf{1}\{\text{siblings} > 2\}$);

father's and mother's earnings bins (= 0, < 50K, < 100K, < 200K, < 1M NIS; and parental schooling bins for each parent (12 years; 13–15 years; > 15 years.

B.3 Information on Official Admission Thresholds

The interpretation of the results in this paper relies on institutional details showing that admission chances to computer science and electrical engineering programs do not increase at the 600 cutoff, a finding supported by the empirical analysis presented throughout the paper. While I do not observe official admission requirements for every institution and year, I do observe partial records for specific years and universities. In all available cases, admission thresholds for both programs are substantially above 600, making it nearly impossible to gain admission at this level. For example, in 2012 a candidate with a score of 600 required a matriculation GPA of 114 to enter the computer science program at Tel Aviv University, which is exceptionally high and rare.[23] Similar requirements are observed in other universities and years, both for computer science and for electrical engineering.

Some programs may also impose additional subject-specific requirements beyond the Sechem index, such as minimum UPET or matriculation scores in mathematics or science. For example, admission to computer science or electrical engineering often requires participation in four- or five-credit matriculation programs in mathematics or related subjects, with a minimum score threshold. At Ben Gurion University in 2013, all scientific programs—including computer science and electrical engineering—required a UPET score of at least 600 in addition to the Sechem requirement. In practice, however, this requirement was not binding for computer science or electrical engineering applicants, given that their Sechem thresholds were substantially higher.

There are also indirect pathways into these programs. Israeli universities operate pre-academic programs designed to improve students' admission chances, and students may transfer into computer science or electrical engineering after completing their first year in another program.

Unfortunately, I lack comparable information on college applications and admission decisions during the study period. However, it is important to note that college admission requirements are typically lower than those at universities. Thus, it is likely that a score of 600 could suffice for enrollment in college-level computer science or electrical engineering programs at least in specific colleges.

[23] These values can exceed 100 because of the GPA calculation method, which adds 10–30 bonus points for certain advanced high school tracks (e.g., five-credit programs).

Appendix C Heuristic Self-Evaluation among Top Scorers

The main analysis centers on the 600-point threshold, where the decision to pursue selective STEM degrees versus less rewarding alternatives is most relevant. I also examine outcomes just above 700, the highest round score in this setting. Only about 5 percent of Israeli test-takers cross this level, which limits the sample size and narrows the analysis to top achievers. In terms of admissions, students scoring around 700 typically gain direct entry into nearly all degree programs, with the exception of medical school, which usually requires higher scores.

Table A.24 reports regression discontinuity estimates for scoring above 700 on the first test. Panel A presents application and retake outcomes. In the short term, there is a small and insignificant increase in applications to lucrative STEM programs, though the baseline rate is already high at 40 percent. When splitting the sample by quantitative advantage, the effect differs sharply: a 9.6-percentage-point increase from a 49 percent baseline among students with an advantage, versus a 1.7-percentage-point decline from a 32 percent baseline among those without.

Panel B presents degree enrollment results. There is a marginally significant rise in ever enrolling in any degree, from a baseline of 97 percent. More strikingly, enrollment in advanced degrees increases significantly by 8 percentage points, from a 49 percent baseline. Using an indicator for PhD enrollment, results also show a significant increase of 2.4 percentage points, relative to a baseline of 9.2 percent.

Panel C examines labor-market outcomes. Results show a shift of 2.2–2.5 percentage points from self-employment to salaried employment, with no significant effects on either tech employment or wages.

These results indicate that left-digit bias influences students' decisions in Israel not only at the 600 cutoff but also at higher thresholds. A plausible interpretation is that the perceived boost in success encouraged students to invest more in their studies and increased their likelihood of pursuing advanced degrees.

Appendix D Returns Relative to Observational Differences

This appendix section formalizes and presents results for the local wage returns relative to observational differences between compliers' chosen degrees. I follow the definitions in the main text. s_i denotes the first test score. The instrument is $Z_i \equiv \mathbf{1}\{s_i \geq 600\}$. $D_i^*(z) \in \{d_1, \ldots, d_K\}$ denotes the potential degree program (defined by field and institution type) chosen by student i if $Z_i = z$, where K is the number of available programs. $Y_i(d)$ denotes the potential outcome (e.g., wage at ages 30–32) if $D_i^* = d$.

I assume continuity and the exclusion restriction as detailed in the main text, and transform the treatment to an ordered continuous version using the observed average wage associated with the chosen degree. Formally, for each degree program d (defined by field and institution type), compute the mean wage outside the estimation window: $m(d) = \mathbb{E}[Y_i \mid D_i^* = d,\ s_i \notin [581, 619]]$. For each individual in the regression discontinuity sample, assign $m_i = m(D_i^*)$, the observed program-specific mean corresponding to the chosen degree. Define the potential ordered treatment values as $m_i(z) = m(D_i^*(z)), \quad z \in \{0, 1\}$. Positive monotonicity with respect to m requires: $\Pr(m_i(1) \geq m_i(0)) = 1, \qquad \Pr(m_i(1) > m_i(0)) > 0$. The first condition rules out defiers in the ordered treatment, and the second ensures the existence of a positive mass of compliers.

Two caveats to this assumption are worth noting. First, students' potential wages across fields may be heterogeneous, so individual rankings of programs need not coincide with the observational ranking $m(d)$. Some students may, in principle, react to the round score by choosing a program with a lower $m(d)$ but a higher idiosyncratic payoff. Empirically, however, the effects concentrate in a narrow set of high-earning STEM degrees, consistent with the dominance of the tech sector in Israel.

Second, the instrument may reduce retesting effort for some students, potentially shifting them into less rewarding programs and thereby violating monotonicity. While there is evidence of reduced retesting in the main sample, it does not appear to translate into meaningful differences in degree choice, so it is unclear whether monotonicity is truly violated. Moreover, this reduction is concentrated among non–quantitatively advantaged students. Accordingly, I also report estimates for quantitatively advantaged students, who do not reduce retesting and for whom the monotonicity condition is more plausible.

To assess monotonicity, I implement stochastic dominance tests. Figure A.9 plots the marginal distributions of degree-associated wages above and below the 600 cutoff, estimated

via regression discontinuity, and indicates that the distribution above 600 weakly dominates that below.

Then, I estimate the following fuzzy regression discontinuity:

$$\delta^m = \frac{\lim_{s\downarrow 600} \mathbb{E}[Y_i \mid s_i = s] - \lim_{s\uparrow 600} \mathbb{E}[Y_i \mid s_i = s]}{\lim_{s\downarrow 600} \mathbb{E}[m_i \mid s_i = s] - \lim_{s\uparrow 600} \mathbb{E}[m_i \mid s_i = s]} \tag{A1}$$

To interpret δ^m, define individual-level changes under the instrument as $\Delta_i^Y = Y_i(D_i(1)) - Y_i(D_i(0))$ and $\Delta_i^m = m_i(1) - m_i(0)$. By definition, and by the exclusion restriction, $\Delta_i^Y = \Delta_i^m = 0$ for stayers, that is, when $D_i(1) = D_i(0)$.

Let C denote the set of compliers (those with $\Delta_i^m > 0$) and let $p_C = \Pr(i \in C)$ denote their share. Under continuity, exclusion, and monotonicity in m, the reduced form and first stage can be expressed as $\lim_{s\downarrow 600} \mathbb{E}[Y_i \mid s_i = s] - \lim_{s\uparrow 600} \mathbb{E}[Y_i \mid s_i = s] = p_C\, \mathbb{E}[\Delta_i^Y \mid C]$, and $\lim_{s\downarrow 600} \mathbb{E}[m_i \mid s_i = s] - \lim_{s\uparrow 600} \mathbb{E}[m_i \mid s_i = s] = p_C\, \mathbb{E}[\Delta_i^m \mid C]$. It follows that $\delta^m = \frac{\mathbb{E}[\Delta_i^Y|C]}{\mathbb{E}[\Delta_i^m|C]} = 1 + \frac{\mathbb{E}[\Delta_i^Y - \Delta_i^m|C]}{\mathbb{E}[\Delta_i^m|C]}$.

This expression makes the benchmark and deviations transparent. If, for local compliers at the cutoff, the average wage gain from switching programs equals the average change in the program-level observed mean wage $m(d)$ associated with their chosen degrees, then $\mathbb{E}[\Delta_i^Y \mid C] = \mathbb{E}[\Delta_i^m \mid C] \quad \Rightarrow \quad \delta^m = 1$. Estimates with $\delta^m > 1$ indicate that complier gains exceed the average observational differences, while estimates with $\delta^m < 1$ indicate that complier gains fall short of these differences.

There are two main reasons why δ^m may differ from one. First, $m(d) = \mathbb{E}[Y \mid D = d]$ is not a causal parameter. It reflects a mix of causal returns to program d and selection on unobservables. By contrast, the regression discontinuity design eliminates this selection bias by comparing students who are locally similar in ability and background but differ in their program choice only because of the round-score instrument. As a result, the regression discontinuity isolates causal effects, while $m(d)$ conflates causal effects with selection. Second, treatment effects may be heterogeneous across individuals. Even if $m(d)$ were purged of selection, it would still capture an average return across all students in program d, not necessarily the return for those marginal students who switch programs at the cutoff. Compliers may differ systematically from inframarginal students in ways that make their returns larger or smaller than the program average.

Because selection is plausibly positive into high-return STEM degrees, the use $m(d)$ will typically overstate the true causal return to those programs. Using $m(d)$ as the treatment variable therefore tends to bias the denominator of δ^m upward, which in turn biases δ^m

downward relative to the benchmark where program effects are measured causally. Thus, observing $\delta^m > 1$ despite this downward bias is particularly informative: it implies that complier returns exceed even the inflated observational differences, pointing to especially high gains for marginal students who enter lucrative STEM fields at the cutoff.

Moreover, to probe sensitivity to selection, I also construct a control-adjusted program index and re-estimate δ^m using this alternative treatment measure. Let $m'(d)$ denote the program-specific component from a saturated wage regression estimated outside the regression discontinuity window:

$$Y \;=\; \sum_d \alpha_d \, \mathbf{1}\{D = d\} \;+\; g(X) \;+\; u, \qquad s_i \notin [581, 619],$$

where $g(X)$ flexibly controls for predetermined characteristics. These include decile indicators for first and maximum scores in each of the three domains of the university entrance test; total number of tests; deciles of Bagrut GPA; total Bagrut credits and the number of credits in each scientific subject (mathematics, physics, biology, chemistry, computer science); demographics (sex, birth year, population group); parental earnings terciles for each parent at ages 14–16; and parental schooling terciles for each parent. Missing values are accounted for using explicit "missing" categories.

Table A.25 reports the estimates of δ^m using both m and m' as the treatment. In the full sample, the results are difficult to interpret because of a noisy first stage. The estimates for quantitatively advantaged students reveal sizable returns for compliers. Using the raw observational mean as the treatment, the coefficient is 1.6 when wages are measured in levels and 0.7 when using log wages. When the control-adjusted index m' is used, the coefficient rises above one in both cases (2.7 in levels and 1.9 in logs). This pattern is consistent with positive selection bias in the observational measure: the unadjusted m incorporates both program effects and selection, which tends to push δ^m downward. Adjusting for observables removes part of this bias, and the resulting estimates suggest that compliers in this setting earn more than the average student entering the same programs. Specifically, the lower bounds of the 95% confidence intervals still imply large returns relative to program averages (1.35 in levels and 0.75 in logs). Given that even these estimates are likely biased downward relative to the true heterogeneity in complier returns, the evidence strongly supports the conclusion that compliers do not earn less—and may in fact earn more—than average students in the same programs.

Appendix E Binarization

In this appendix I examine the conditions under which binarizing the treatment allows for a valid fuzzy regression discontinuity design in our setting, and I propose alternative estimates for cases where these conditions may be violated. I follow the definitions from the main text and Appendix D. Let s_i denote the first test score, and define the instrument as $Z_i \equiv \mathbf{1}\{s_i \geq 600\}$. The treatment is an indicator for pursuing a lucrative STEM degree, obtained by binarizing an underlying multivalued treatment that orders programs by their associated wages: $D_i = \mathbf{1}\{m_(D_i^*) > 17.5\}$. Finally, let $Y_i(d)$ denote the potential outcome (e.g., wages at ages 30–32) if $D_i = d$.

Note that under these definitions, with monotonicity in m, and without assuming that the binarization is valid, individuals can be classified into three mutually exclusive complier sets:

$$C_s = \{m_i(1) > m_i(0) > 17.5\}, \qquad C = \{m_i(1) \geq 17.5 > m_i(0)\}, \qquad C_n = \{17.5 > m_i(1) > m_i(0)\}.$$

Only the middle group C corresponds to the switch from a non-STEM degree to a lucrative STEM degree. The groups C_s and C_n are always- and never-takers, respectively, with respect to the binary treatment of pursuing STEM (D).

The reduced-form (RF) effect of the instrument on the outcome can then be written as:

$$\begin{aligned} &\underbrace{\lim_{s\downarrow 600} \mathbb{E}[Y_i \mid s_i = s] \;-\; \lim_{s\uparrow 600} \mathbb{E}[Y_i \mid s_i = s]}_{\text{RF}} \\ &\quad = \mathbb{E}[\Delta_i^Y \mid C]\Pr(C) \;+\; \underbrace{\mathbb{E}[\Delta_i^Y \mid C_n]\Pr(C_n) \;+\; \mathbb{E}[\Delta_i^Y \mid C_s]\Pr(C_s)}_{=:B}, \end{aligned} \tag{A2}$$

where B collects bias terms arising from movements within the lucrative STEM or non-STEM groups.

The first stage with respect to D is

$$\underbrace{\lim_{s\downarrow 600} \mathbb{E}[D_i \mid s_i = s] - \lim_{s\uparrow 600} \mathbb{E}[D_i \mid s_i = s]}_{\text{FS}} = \Pr(C).$$

Hence,

$$\widehat{\delta} = \frac{\text{RF}}{\text{FS}} = \underbrace{\mathbb{E}[\Delta_i^Y \mid C]}_{\text{target LATE}} + \frac{B}{\Pr(C)}.$$

The bias term B disappears under either of two scenarios (see also Andresen and Huber, 2021): (i) there are no within-group switches across the threshold ($\Pr(C_n) = \Pr(C_s) = 0$), meaning all compliers move from non-STEM to high-earning STEM degrees; or (ii) such within-group switches have no effect on earnings ($\mathbb{E}[\Delta_i^Y \mid C_n] = \mathbb{E}[\Delta_i^Y \mid C_s] = 0$), which holds if movements between programs within the same category do not affect wages.

Diagnostics. I implement a diagnostic test proposed by Andresen and Huber (2021) for instrumental variable settings where a multi-valued treatment is collapsed into a binary indicator. A key implication of the conditions for valid binarization is that, if we estimate first-stage regressions using different cutoff values, the estimated first stage should increase up to the chosen cutoff (17.5 in our case) and then decrease. Figure A.9 plots the marginal distribution of degree-associated wages above and below the 600 cutoff, based on a series of regression discontinuity estimations, and shows evidence that this holds in our setting. In later versions of the paper, I will report the formal test, too. The figure shows a monotonic increase in the share of students pursuing degrees above the threshold, with a clear peak at 17.5. This corresponds to roughly 4 percent of students in the main sample moving from non-high-earning to high-earning degrees, and about 10 percent among quantitatively advantaged students.

Nevertheless, as pointed out by Andresen and Huber (2021), this condition is necessary but not sufficient. We cannot entirely rule out more complex patterns of response. To address this concern, I also consider bounds that require weaker restrictions relative to the binary fuzzy regression discontinuity.

Bounds. To construct bounds, I first note that the assumption $\Pr(C_s) = 0$ is relatively weak in our setting, as the lucrative STEM category comprises only a handful of very rewarding degrees. Moreover, the data provides support for this restriction. As shown in Figure A.9, the share of students who pursue the very top-earning programs is extremely small, and compliers never reach them. For instance, the first stage for $m \geq 22$ is exactly zero (both in the full and quantitatively advantaged samples), which implies that no student crosses into any degree with associated wages above 22.

xxx define xxx Formally, let $\beta(j)$ measures the probability mass of individuals whose program index crosses j when the instrument shifts from $Z = 0$ to $Z = 1$. If $j > 17.5$ and $\beta(j) = 0$, then no unit crosses j ($\Pr\big(m_i(0) < j \leq m_i(1)\big) = 0$), and therefore no high-group movers (C_s) can exist in the region $[j, \infty)$. By monotonicity, this rules out violations above that threshold. The same argument applies symmetrically for $j < 17.5$, where $\beta(j) = 0$ rules out low-group movers (C_n) in the region $(-\infty, j]$.

In our data, $\beta(j) = 0$ for all $j \geq 22$, which ensures the absence of high-group movers into degrees above that point. This substantially limits the scope for within-STEM violations. By contrast, movements within the non-STEM group remain more plausible. For example, crossing the 600 threshold can shift students from sociology to economics—a transition that increases earnings but leaves the STEM indicator unchanged. Such movements violate the exclusion restriction, since the instrument affects wages not only by changing the binary STEM status D, but also by reallocating students across other fields. This generates upward bias: the reduced form captures the wage gains of both binary compliers and never-STEM movers, while the first stage reflects only the former.

We maintain $\Pr(C_s) = 0$ and allow $\Pr(C_n) > 0$. In a fuzzy regression discontinuity, the treated complier mean is point-identified using the modified outcome $\widetilde{Y}_i = Y_i D_i$:

$$\Big(\lim_{s\downarrow 600} \mathbb{E}[YD \mid s] - \lim_{s\uparrow 600} \mathbb{E}[YD \mid s]\Big) = \Pr(C)\,\mathbb{E}[Y(D(1)) \mid C] \;+\; \Pr(C_s)\,\mathbb{E}[\Delta^Y \mid C_s],$$

Never-STEM compliers do not contribute because $D = 0$ for them on both sides; hence the fuzzy Wald ratio with outcome YD and treatment D equals

$$\frac{\lim_{s\downarrow 600} \mathbb{E}[YD \mid s] - \lim_{s\uparrow 600} \mathbb{E}[YD \mid s]}{\lim_{s\downarrow 600} \mathbb{E}[D \mid s] - \lim_{s\uparrow 600} \mathbb{E}[D \mid s]} = \mathbb{E}[Y(D(1)) \mid C] \;+\; \frac{\Pr(C_s)}{\Pr(C)}\,\mathbb{E}[\Delta^Y \mid C_s].$$

Under $\Pr(C_s) = 0$, this equals $\mathbb{E}[Y(D(1)) \mid C]$.

Our target is $\mathbb{E}[Y(D(1)) - Y(D(0)) \mid C]$. Given $\mathbb{E}[Y(D(1)) \mid C]$, it suffices to recover $\mathbb{E}[Y(D(0)) \mid C]$. We get that from the outcome distribution of untreated individuals just below the cutoff. By the fuzzy regression discontinuity assumptions, students below 600 who do not pursue STEM are a mixture of compliers (who would switch at 600) and never-takers. Formally, define

$$p_{0-} \;:=\; \lim_{s\uparrow 600} \Pr(D = 0 \mid s), \qquad p_{0+} \;:=\; \lim_{s\downarrow 600} \Pr(D = 0 \mid s).$$

Monotonicity implies $p_{0+} = \Pr(D_i(0) = 0,\ D_i(1) = 0)$, that is, the probability of being a never-taker, and $p_{0-} = \Pr(D_i(0) = 0,\ D_i(1) = 0) + \Pr(C)$, that is the probability of being either a never-taker or complier.

Let $p_C = \Pr(C)$ which is identified by the first stage. The complier share among units with $D_i = 0$ just below the cutoff is then $\omega_C = \frac{\Pr(C)}{\Pr(D_i(0)=0,\, D_i(1)=0)+\Pr(C)} = \frac{p_C}{p_{0-}}$.

Now, we can obtain conservative estimates for the untreated complier mean. A worst-case scenario with no additional assumptions is given by trimming the untreated distribution so that compliers are assumed to be those with the highest wages (Lee, 2009). Formally, let

$Q^{-}_{Y|D=0}(q)$ denote the left-limit ($s \uparrow 600$) q-quantile of Y within $D = 0$. Under the extremal ordering that places all compliers weakly *above* never-takers in the $Y(0)$ distribution,

$$\overline{\mu}_0^C \;:=\; \mathbb{E}\big[Y \;\big|\; s \uparrow 600,\; D = 0,\; Y > Q^{-}_{Y|D=0}(1-\omega_C)\big]$$

is an upper bound on $\mathbb{E}[Y(0) \mid i \in C]$. Therefore, given that $\overline{\mu}_0^C \;\geq\; \mathbb{E}[Y(0) \mid C]$, we get $\mathbb{E}[Y(1) \mid C] - \overline{\mu}_0^C \;\leq \mathbb{E}[Y(1) - Y(0) \mid C]$.

This is a worst-case scenario, which is highly restrictive and therefore yields an insignificant lower bound for the treatment effect. I focus on monthly wages at ages 30–32. For simplicity, I implement by calculating averages within a narrow 10-point window around the 600 cutoff. As shown earlier, this approach produces results very similar to the main local linear regression discontinuity specification. Using this conservative method, I estimate worst-case bounds of -4,000 in the main sample and positive but insignificant of 1,200 NIS in the quantitative advantaged subsample, with a bootstrap standard error of 1,700 NIS.

I next propose a tighter yet conservative trimming approach, which relies on the program index m rather than actual wages. Define $Q^{-}_{m|D=0}(q)$ as the left-limit quantile of m among untreated units ($D = 0$). Under the worst-case ordering, where all compliers lie strictly above never-takers in the $m(0)$-distribution,

$$\overline{\mu}^C_{0,m} \;:=\; \mathbb{E}\big[Y \;\big|\; s \uparrow 600,\; D = 0,\; m > Q^{-}_{m|D=0}(1-\omega_C)\big] \quad \text{is an upper bound on } \; \mathbb{E}[Y(0) \mid C],$$

This approach may not provide an upper bound for untreated complier mean in all cases. For example, compliers could have lower counterfactuals yet be especially talented and thus earn much more relative to the average of their degrees—the following assumption justifies interpreting it as a worst-case scenario bound: Let $\mu_0(t) := \mathbb{E}\big[Y(0) \mid m(0) = t,\; s \uparrow 600\big]$. Then $\mu_0(t)$ is nondecreasing in t. In any case, this constitutes a conservative approach that permits cautious inference on local returns in our setting, while also allowing for violations of the exclusion restriction.

Figure A.10 illustrates this approach for the subsample of quantitatively advantaged students. The figure shows the wage distribution of untreated individuals below the 600 threshold and highlights the estimated treated-complier mean (31,200 NIS). The untreated group consists of both untreated compliers—the target counterfactual—and never-takers. If these two groups are similar, the counterfactual mean is simply the group average (14,200 NIS), yielding local returns of 17,000 NIS.

Using the estimated complier and never-taker shares (8.9 and 59 percent) I derive bounds for the untreated-complier mean by trimming the distribution of untreated outcomes just below

the cutoff, keeping the top $\frac{8.9}{8.9+59} \approx 13$ percent of the sample. This gives a counterfactual mean of 30,000 NIS and an estimated return of 1,200 NIS (bootstrapped standard error: 1,700).

I then repeat the trimming exercise using degree-based ranking. The mean wage among the 13% of students in the highest-wage degrees is 19,800 NIS. This remains well below the treated-complier mean, implying local gains of at least 12,500 NIS (standard error: 1,100).