Heuristic Self-Evaluation in High-Stakes Decisions:
Fields of Study and Long-Term Earnings

Yoav Goldstein*

27 October 2025

Click here for the most recent version

Abstract

This paper examines how uncertainty about one's ability interacts with cognitive heuristics to shape high-stakes educational choices and career paths. Using administrative data, I implement a regression discontinuity at a round-number threshold on the national university entrance test. Crossing the threshold on the first test sharply increases the likelihood of pursuing high-return STEM degrees, even though it confers no formal admission advantage and often falls below the cutoff for these programs. This behavior is consistent with left-digit bias: students interpret the score's left digit as a meaningful signal, leading them to improve their scores through retesting, apply to, and ultimately enroll in STEM programs. In subsequent years, they are more likely to work in the tech sector, earn higher wages, and reach the top of the income distribution. Leveraging this behavioral discontinuity, I estimate large returns to STEM degrees among students uncertain about applying. The findings suggest that heuristic self-evaluation can generate persistent disparities in economic opportunity that are unrelated to actual ability.

JEL Codes: I26, J24, J31, D91

*The Eitan Berglas School of Economics, Tel Aviv University, yoavg2@mail.tau.ac.il. 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". Research was conducted in the research room of the Central Bureau of Statistics based on de-identified individual record files prepared by the Central Bureau of Statistics.

1 Introduction

Human capital investments shape long-term life outcomes, yet students often face uncertainty about their own ability when making such investments (Altonji, 1993). Selective educational programs use test scores, which are noisy signals of ability, to allocate admission opportunities. This paper examines how uncertainty and cognitive heuristics influence educational choices in test-based admission systems and investigates their long-term consequences. It focuses on a high-stakes decision: whether to invest in technology-oriented education in a context of high demand for technological skills.

I exploit left-digit bias, the tendency to focus on the leftmost digit of a number while partially ignoring other digits (Poltrock and Schwartz, 1984; Korvorst and Damian, 2008). This bias generates quasi-experimental variation in perceived test performance around round-number thresholds. Accordingly, I implement a regression discontinuity design around a round-number threshold on students' university entrance test. The threshold does not affect admission prospects, yet students appear to interpret it as a meaningful signal. Crossing it on the first test sharply increases applications to high-return STEM programs. Students seem aware that their initial scores were insufficient: many improve them by retaking the test and ultimately enroll in high-return STEM.

This raises the question of whether these students actually realize gains, or if the increase in STEM enrollment reflects overconfidence. Evidence indicates actual gains: years later, they are more likely to work in the technology sector, earn higher wages, and rank among the top earners. Together, the results show that uncertainty about one's own ability, combined with heuristic interpretations of noisy scores, can deter capable students from pursuing high-return degrees and lead them to forgo significant economic opportunities.

As detailed in Section 2, this study uses administrative data from the Israeli Central Bureau of Statistics. The dataset links students' entrance test scores to 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 employment and earnings for up to two decades after the initial test.

In Israel, technology-oriented careers are associated with much higher wages. Among university graduates, degrees in selective STEM fields (computer science and electrical engineering) lead to high employment rates in the technology sector and wages about twice as high as those in most other fields. By their early thirties, the median graduate of these

[1]Because of mandatory military service, about half of first-time test-takers in Israel are aged 21 or older. These test-takers usually apply soon after taking their first test, leaving little room for behavioral responses.

programs earns more than 90 percent of workers nationwide, including those at all ages and experience levels.

Admission to Israeli universities is determined by a composite of the entrance test score and matriculation GPA, based on each applicant's highest score in both components. Decisions rely solely 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 in this context.

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, as 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 round-number threshold. This is plausible given the institutional setting, where score manipulation is highly unlikely, and is supported empirically by the smoothness of score distributions and baseline characteristics around the threshold.

Regression discontinuity estimates presented in Section 4 show that crossing the round-number threshold increases the likelihood of applying to selective STEM degrees within five years of the test by 3.8 percentage points from a baseline of 9.6 percent.[2] The effect is persistent and reflects a shift in field choice rather than an increase in total applications or institutional selectivity. It is not driven by a decrease in applications to other high-return STEM degrees. The overall likelihood of applying to any high-return STEM program—including computer science, electrical engineering, and other related degrees—increases by 4.8 percentage points.

These results are consistent with heuristic self-evaluation influencing decisions to apply to selective STEM degrees: students scoring just above a round number interpret their performance as a stronger signal and aim higher. I directly test alternative explanations. Admission probabilities are smooth at the threshold, and the pattern does not align with the admission-misperception channel, in which applicants mistakenly treat the round number as a formal cutoff. Most STEM applicants retest and substantially raise their scores, suggesting awareness that their initial scores were insufficient for admission.

Administrative data, however, do not allow direct identification of the specific psychological mechanism. The response could reflect an internal shift in self-perceived ability, a reaction to how peers interpret the score, or the use of round numbers as reference points for social comparison. While these channels cannot be disentangled empirically in this setting, I test

[2]The appendix reports results for the next higher round-number threshold as well. For those scoring around this threshold in the first test, application rates are much higher, and the left-digit bias manifests instead in behaviors such as pursuing advanced degrees, including in STEM.

for social influence by examining siblings' behavior: when an older sibling scores just above the round number, younger siblings become more likely to take the test in subsequent years.

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 comprises 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 relatively higher quantitative scores are most likely to increase applications to selective STEM degrees. By contrast, students without a quantitative advantage respond differently: they are less likely to retake the test. This pattern may reflect a distinct but related behavioral response to round numbers, treating them as goals (Pope and Simonsohn, 2011).

Examining degree enrollment, I find that about half of the marginal applicants ultimately enroll in selective STEM programs. Among those not enrolled, many shift to other high-earning STEM degrees. Overall, crossing the threshold increases the probability of enrolling in one of Israel's most lucrative degree programs, all in STEM fields, by 3.8 percentage points relative to a baseline of 22 percent.

Tracking this sample into the labor market reveals substantial gains. By 2022–2023, at ages 30-43, those scoring just above the threshold are 3.5 percentage points more likely to work in the technology sector. Their annual earnings are higher by 13,200 NIS (approximately 3,500 USD, or 5 percent) and their monthly wages by 1,400 NIS (approximately 400 USD, or 4 percent). These effects emerge roughly ten years after the test and persist for at least another decade, consistent with sustained returns to human capital investment. Those scoring above the threshold are also more likely to reach the top of the earnings distribution: the probability of being in the top 5 percent rises by 1.6 percentage points from a baseline of 6.8 percent.

Section 5 builds on these results—that scoring above the round number increases both STEM degree enrollment and long-term earnings—to estimate the causal returns to STEM in this setting. Unlike conventional regression discontinuity designs exploiting formal admission cutoffs (e.g., Bleemer and Mehta, 2022), this analysis leverages quasi-random variation in STEM enrollment induced by changes in application behavior. Identification relies on a strong exclusion restriction: crossing the round number must affect outcomes only through the decision to pursue a STEM degree. I focus on the subsample with strong quantitative preparation, where violations are less likely, and provide evidence supporting this assumption. The estimates indicate large returns in monthly wages by ages 30–32, of about 20,000 NIS (roughly 120 percent). I also offer alternative estimates that do not rely on the exclusion restriction and still imply substantial gains.

These findings are relevant for education systems that place heavy weight on admission tests. Policy debates over their use have intensified in recent years. Much of the research examines whether such tests predict later outcomes, with mixed results (Rothstein, 2004; Chetty et al., 2023; Friedman et al., 2025). Other work shows that, in some settings, admission tests can promote mobility (Chetty et al., 2023; Sacerdote et al., 2025). Focusing on students' decision-making, research also shows that complexity and opacity in admissions can distort enrollment choices (Dynarski et al., 2021; Kapor, 2024).

This study's findings highlight a distinct limitation that arises even in transparent systems where admission depends solely on test scores. Students may interpret noisy signals heuristically, leading to distorted educational choices with long-term consequences. These results caution against heavy reliance on test scores and point to scope for interventions within existing systems—for example, redesigning score reports or strengthening the guidance that schools and educators provide at critical decision points.

This paper also relates to studies on returns to education in settings with competitive admissions. A longstanding body of work estimates the returns to selective programs among marginally admitted students (Hoekstra, 2009; Zimmerman, 2014; Anelli, 2020; Mountjoy, 2024), with growing evidence that the field of study is a key determinant of returns (Hastings et al., 2013; Kirkeboen et al., 2016; Heinesen et al., 2024; Bleemer and Mehta, 2022). Another central question in this literature concerns the returns to attending selective programs for less-prepared students. The mismatch hypothesis posits that less-prepared students may be harmed in selective environments because they are less likely to graduate and capture the associated benefits—a central argument against affirmative action policies (Sowell, 1978).

Focusing on STEM, my findings instead align with evidence from Colombia (Ng and Riehl, 2024), where less-prepared students benefited more from selective STEM degrees. In my setting, lower-scoring students who later pursue STEM earn returns at least as large as the average. These results add to recent U.S. evidence (Bleemer, 2021, 2022; Black et al., 2023) challenging the mismatch hypothesis.

A key distinction is that I exploit quasi-random changes in enrollment arising from application behavior. The policy implications differ: whereas admission-margin estimates inform the benefits of expanding program capacity, the results here show that policies can also increase impact by encouraging hesitant students to apply. Such encouragement complements capacity expansion by broadening the applicant pool with strong candidates who might otherwise self-select out. Still, the scope of such policies appears limited in reducing enrollment gaps, as effects are concentrated among students with quantitative preparation, making it unlikely to achieve large reductions in gender or socioeconomic disparities.

Additionally, this paper contributes to the literature on human capital decisions. Prior work shows that uncertainty about one's own ability shapes educational choices (Altonji, 1993; Zafar, 2011; Arcidiacono et al., 2012; Stinebrickner and Stinebrickner, 2012; Wiswall and Zafar, 2015; Papay et al., 2016; Avery et al., 2018; Hakimov et al., 2023; Li and Xia, 2024; Arcidiacono et al., 2025).[3] The results here contribute in three key ways. First, they suggest that heuristics create behavioral frictions in the process of learning about one's own ability. Second, they document under-evaluation in field-of-study choice. Unlike prior work emphasizing overconfidence and subsequent attrition from STEM (Stinebrickner and Stinebrickner, 2014), I show that students forgo STEM despite having the ability to complete these programs successfully. Third, following students for two decades reveals substantial long-run economic losses from this underevaluation.

Finally, this paper contributes to the literature on heuristics and biases. Round number heuristics and left-digit bias have been documented in many domains, including consumer demand and firm pricing (Lacetera et al., 2012; Strulov-Shlain, 2023), wage setting (Dube et al., 2025), expert judgment (Olenski et al., 2020; Shurtz, 2022), and goal setting (Pope and Simonsohn, 2011). Similar behavior appears in SAT retaking in the United States (Pope and Simonsohn, 2011; Goodman et al., 2020). I document similar effects, but I also emphasize an additional channel: self-evaluation. Evidence from China also suggests a similar mechanism, as applicants become more ambitious in their college applications after crossing round numbers in the Gaokao (Li and Qiu, 2023).

While previous studies document left-digit bias across various domains, including responses to test scores, this paper makes several distinct contributions. First, by focusing on a specific decision, it allows direct testing of alternative explanations—such as formal admission advantage or admission misperceptions—providing evidence against them. Second, it demonstrates that heuristics influence even one of the most consequential choices for young adults in Israel. Third, it links these behaviors to long-term outcomes, documenting substantial and persistent earnings gains. Together, the results demonstrate that heuristics and biases—often regarded as minor, arising from inattention, and unlikely to persist when stakes are high—play a significant role in high-stakes decisions with lasting consequences.

[3] Test score shocks (e.g., Ebenstein et al., 2016; Lavy and Megalokonomou, 2019; Landaud et al., 2024) and within-class rank (e.g., Diamond and Persson, 2016; Elsner and Isphording, 2017; Murphy and Weinhardt, 2020; Elsner et al., 2021; Denning et al., 2023) have also been shown to affect choices and outcomes, possibly through related mechanisms.

2 Institutional Background and Data

Israel is a high-income OECD economy with a highly developed technological sector. The OECD (2023) describes Israel's "vibrant high-tech industry", which drives productivity and export performance but also contributes to a dual economy with large productivity and wage gaps between sectors. According to Israel Innovation Authority (2021), the high-tech sector accounts for about 10 percent of total employment, around 15 percent of GDP, roughly half of exports, and about one-quarter of personal income tax revenues.

This paper examines decisions to pursue high-return, technology-oriented degrees. Given the success of Israel's tech sector, these are high-stakes choices with long-term implications for economic opportunities. This section provides institutional background on Israel's postsecondary education system and admissions, describes the administrative data used in the analysis, presents evidence on the labor-market dominance of STEM graduates, and elaborates on STEM admissions and the definition of the study sample.

2.1 Postsecondary Education Institutions and Admissions

Israel's postsecondary education system is organized as a dual structure comprising public universities and academic colleges. Universities are the most established, selective, and research-oriented institutions. They dominate in STEM fields, medicine, and graduate education, and award nearly all doctoral degrees. In contrast, public and private colleges are generally considered lower-tier institutions, as reflected in their admission standards, teaching focus, and graduates' labor-market outcomes (e.g., Achdut et al., 2019).

There is also heterogeneity across institutions within the university sector. According to the Central Bureau of Statistics (CBS), three institutions—the Hebrew University of Jerusalem, Tel Aviv University, and the Technion—are regarded as Israel's elite universities. These institutions are generally more selective, though admission standards vary substantially across fields of study within each university. The system is overseen by the Council for Higher Education, the statutory body responsible for accreditation, funding, and quality assurance. All universities, and some colleges, receive substantial state funding, and tuition levels are low by international standards, averaging about 3,000 USD per year.[4]

Admissions are decentralized: each institution sets its own criteria and manages its own applications. Applicants may apply to multiple institutions and rank up to three fields of study per institution. Admission to postsecondary programs is determined mechanically by a

[4]For comparison, the average annual tuition at public colleges in the United States is about 10,000 USD (source: https://nces.ed.gov/fastfacts/display.asp?id=76, retrieved on October 15th, 2025)

combination of University Psychometric Entrance Test (UPET) scores and GPA from Israel's Bagrut (high school matriculation) exams.[5] During the sample period, admissions did not involve essays, recommendation letters, or extracurricular activities.

Degree programs, defined as a specific field at a given institution, set annual admission thresholds, typically based on a weighted average of applicants' highest UPET score and Bagrut GPA. Thresholds vary across programs and years and are generally published online after each admission cycle.

The UPET is a standardized test comprising quantitative, verbal, and English components. Each section is scored from 50 to 150, and the composite total is standardized to a 200–800 scale. The test is administered several times a year across Israel and can be retaken multiple times; universities consider only the highest total score. It is offered in several languages (Hebrew, English, and Arabic) and is relatively inexpensive, with a fixed registration fee of 560 NIS (about 150 USD). A large share of test-takers enroll in commercial preparation courses, which typically cost around 1,000 USD.[6]

2.2 Data Sources and Key Variables

This study uses administrative data from Israel's Central Bureau of Statistics, accessed through its secure research lab. The population includes all students enrolled in Israeli high schools who were in tenth grade between 1995 and 2016. The dataset links multiple administrative sources (detailed in Appendix A), most notably UPET tests and scores, postsecondary education records, and tax data.

Postsecondary education records. The Israeli Council for Higher Education provides data on applications, enrollment, and attainment. Enrollment and attainment records cover

[5] 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 across subjects, with higher proficiency levels receiving 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).

[6] In 2016, 43 percent of course participants reported paying 2,000–5,000 NIS, 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-005056aa4243/2_57aeced1-28b4-ec11-814b-005056aa4246_11_19627.pdf (retrieved September 2, 2025). These figures were likely similar during the study period.

both universities and colleges, whereas application data during our study period are available only for universities. Because high-earning STEM degrees are offered predominantly by universities, this limitation is unlikely to affect the main analysis.

I construct indicator variables for applications (within five years of the first UPET test), and for enrollment and attainment. Applications are restricted to universities, while enrollment and attainment are observed for both universities and colleges. The analysis focuses on two main measures of STEM choice. The first is a narrow definition of "Selective STEM" degrees, including only computer science and electrical engineering programs at universities. The second is a broader empirical definition of "Lucrative STEM" degrees (listed in Table G.1). Focusing first on applications to Selective STEM programs helps isolate underlying mechanisms, since it narrows attention to a small, homogeneous group of highly rewarding degrees and simplifies tests for alternative explanations such as admission advantages. The broader definition captures the overall likelihood of pursuing high-return STEM degrees. Selective STEM is a subset of Lucrative STEM.

Beyond STEM-focused indicators, I also track applications and enrollment across other fields, grouped by institution type and by field groups following the classification of the Israeli Central Bureau of Statistics. Finally, I construct a degree-associated wage measure by assigning each individual the mean wage of graduates from their chosen program at ages 3032.[7] When multiple programs are relevant, I use the maximum value. Based on this measure, I further classify degrees into three wage tiers: low (average wage below 12,500 NIS), mid (12,500–17,500 NIS), and high (above 17,500 NIS). As shown below, the high-wage category consists exclusively of technology-oriented degrees and is therefore used to define the set of Lucrative STEM programs.

The University Psychometric Entrance Test (UPET). The dataset includes all UPET tests for individuals in the sample, provided by the National Institute for Testing and Evaluation. For each test-taker, I retain only the first test for the main analysis, while using subsequent tests to capture retesting behavior and the maximum score achieved. Figure G.1 displays the score distribution.[8]

Labor-market outcomes. The dataset includes individual tax records for all years from 1990 to 2023, excluding 2021 due to missing data. I focus on outcomes in 2022–2023, the most

[7]Individuals in the main regression discontinuity analysis sample are excluded when constructing this outcome.
[8]Because the distribution is approximately normal, very few individuals fall at the extremes. A cell-suppression policy of the Israeli CBS prohibits reporting statistics for very small groups. Accordingly, the analysis is restricted to total scores between 370 and 730.

recent available years, and also present age-specific trajectories and dynamic effects relative to the year of the test. From these records, I construct measures of salaried employment (positive wage income and positive months worked), self-employment (any business income), and employment in the tech sector. "Tech" follows the CBS definition, encompassing ICT services, electronics and communications manufacturing, medical and scientific instruments, pharmaceuticals, aircraft manufacturing, and R&D services (Appendix A).

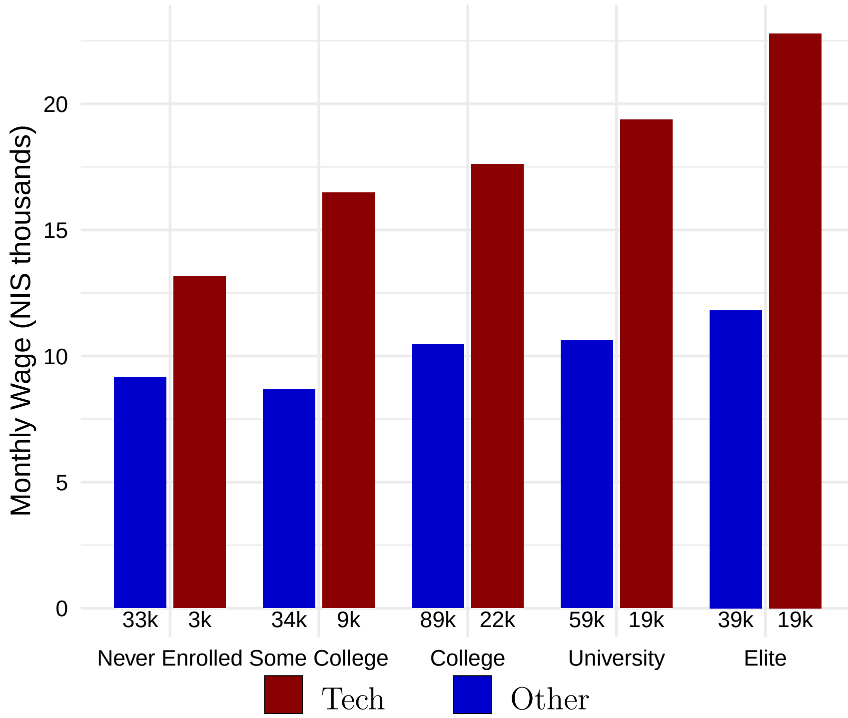
Additional outcomes include monthly wages for salaried workers, annual salaried earnings, and total annual earnings (salaried and self-employment). I also compute age-specific earnings ranks and indicators for belonging to the top 10, 5, or 1 percent of the earnings distribution, defined relative to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

2.3 Labor-Market Outcomes of Technology-Oriented Degrees

Figure 1 illustrates the economic opportunities offered by Israel's tech sector. Across education levels, workers employed in the tech sector earn substantially higher wages. Among graduates of elite universities, average monthly wages in tech are nearly double those outside tech (23,000 versus 12,000 NIS). This gap is much larger than the difference between elite graduates outside tech (12,000 NIS) and individuals who never enrolled in postsecondary education (9,000 NIS). Even individuals without a college degree who work in the tech sector earn more (13,000 NIS) than elite university graduates employed outside tech (12,000 NIS).

As in many other countries, fields of study in Israel differ sharply in their wage premia (Krill et al., 2019; Achdut et al., 2019). The thriving tech sector generates strong demand for STEM skills. Figure 2 illustrates this pattern, showing average wages and tech-employment rates at ages 30–32 for students who have ever enrolled in any undergraduate degree. Each circle represents a field–institution pair. A clear pattern emerges: wages increase with the share employed in tech. Computer science and electrical engineering stand out, exhibiting exceptionally high wages and tech-employment rates across all institution types. Students in these fields earn about 23,000 NIS per month—roughly twice the 11,000 NIS average among non-STEM degrees—and about 60 percent work in the tech sector. Overall, high-wage fields are precisely those with high rates of tech employment.

Figure 1: Average Wage by Postsecondary Education and Sector



Notes: The figure shows labor-market outcomes for students 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 sector group. The total number of observations in each group (education × sector) is shown below each bar. Tech covers technological sectors as defined by the CBS (see Appendix A); "Other" includes all remaining sectors. Elite refers to the three elite universities in the CBS classification (see Appendix A); "Universities" include all other universities. "Some college" refers to students who enrolled in any type of institution but did not graduate.

The figure also illustrates the empirical definition of Lucrative STEM degrees—programs (field–institution pairs) with average monthly wages above 17,500 NIS. As shown in Figure 2, all programs exceeding this wage threshold also have at least 40 percent of students employed in tech. Within this group, nearly all programs are offered by universities (elite or non-elite), with only two exceptions among colleges: computer science and a joint computer science–electrical engineering degree. This reinforces that the lack of college-level application data is not a major limitation. However, analyses of degree enrollment and completion fully capture these programs.

Figure 2: Labor-Market Outcomes of Postsecondary Degree Students

Notes: The figure shows labor-market outcomes for all postsecondary degree students 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 with an average monthly wage above 17,500 NIS. Students in the main analysis group (first-test scores 581-619) are excluded to avoid mechanical overlap with the treatment sample.

Table G.2 complements these patterns by reporting average earnings by sector and degree type, distinguishing between Selective STEM and other degrees. Graduates in computer science and electrical engineering—across all institution types—are far more likely to work in the tech sector than other graduates or non-degree holders, and within tech, they earn substantially higher wages than other workers.

The table also reports parental earnings, providing insight into selection into technology-oriented careers. Across all degree categories, students who work in the tech sector tend to come from higher-income families. At the same time, the sector appears to foster upward mobility: for example, college graduates in computer science or electrical engineering come from lower-income families than graduates of elite non-STEM programs, yet they earn more—23,600 NIS versus 19,500 NIS when employed in tech, and 13,500 versus 11,800 NIS in other sectors.

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 G.2) 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.

Recognizing these substantial wage differentials, Israeli policymakers have sought to expand enrollment in Selective STEM programs. The Council for Higher Education recently launched a national initiative to increase participation in these fields.[9] This initiative parallels international efforts to promote STEM education (e.g., OECD, 2024). Understanding how young adults decide whether to apply to these high-return programs is therefore of direct policy relevance for advanced economies facing persistent demand for highly skilled, tech-oriented workers.

2.4 Admission 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 ever enrolled in Selective STEM degrees, the median admission-relevant UPET score is 685, and the 25th percentile is at 644. These figures illustrate that a score just above 600 is likely insufficient for admission unless paired with an exceptionally strong matriculation GPA. Thresholds for Selective STEM programs are higher than for almost all other programs in Israel, with the exception of medicine (median 738, 25th percentile 721), which generally requires scores above 700.

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

The analysis therefore focuses on the 600-point threshold.[10] 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 may thus be salient due to left-digit bias, yet remains unofficial. Those 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 underpins the key identifying assumption that at the 600 threshold, 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. Section 3 provides empirical evidence that scoring just above 600 has no direct admissions advantage.

2.5 Study Sample

The analysis focuses on first UPET tests taken between 2000 and 2009.[11] 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 before 2000 are excluded because few individuals in the sample took tests earlier. This restriction reduces the sample size by about 10 percent. The main regression discontinuity analysis focuses on students scoring within 20 points of the 600 threshold.

I distinguish three subgroups. First, Arab students are analyzed separately because their background characteristics differ. Among the remaining students, age at first test provides a natural division. Most Israelis begin university between ages 21 and 24, after completing compulsory military service (three years for men, two for women). About half of test-takers sit for the test during or immediately after high school (age 20 or below), while the rest test later. Accordingly, I analyze outcomes separately for three groups: the main sample of younger test-takers (age 20 or below), an older group (age 21 or above), and Arab test-takers.

Table 1 reports summary statistics for each subgroup, both for the full score distribution and for those within the regression discontinuity sample (580–620). Several patterns stand out. Arab students score substantially lower than other test-takers on their first test (426 versus 553 for other young test-takers), which may help explain their lower enrollment in Selective STEM degrees. Comparing younger and older test-takers, initial scores are slightly

[10] 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. Appendix D also documents effects of left-digit bias at the highest round number, 700.

[11] Tests administered between October and December are assigned to the following calendar year, because they could not be used for admission in the same academic year, which typically begins in October in Israel.

lower for the younger group. Older test-takers typically take the test just before applying to university and thus have little time to retest. As a result, their final scores are considerably lower (610 versus 635 within the analysis sample). Accordingly, far fewer older test-takers pursued Selective STEM degrees (2.7 versus 11.1 percent).[12]

Table 1: Summary Statistics by Group

	Main		Older		Arabs	
	(1)	(2)	(3)	(4)	(5)	(6)
Sample	Full	RD	Full	RD	Full	RD
First UPET Total Score	553.37	599.27	563.12	599.71	425.97	597.65
First Quantitative UPET Score	111.93	119.60	111.27	116.93	91.91	123.12
First English UPET Score	110.03	119.24	112.40	120.07	81.85	109.63
First Verbal UPET Score	106.30	114.32	110.09	116.78	84.16	114.88
UPET Retaking (%)	47.10	53.62	22.39	23.73	66.02	83.96
Max UPET Total Score	588.38	635.93	573.82	610.41	463.99	656.61
Parents Postsec. Educ. (%)	49.09	55.25	40.91	44.65	12.35	32.81
Parents High-Income (%)	40.76	46.35	40.18	43.90	8.91	21.54
Female (%)	63.48	60.33	52.87	50.32	61.13	50.94
Selective STEM (%)	9.86	11.09	3.64	2.65	2.46	13.94
Tech (%)	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 UPET scores (total and by domain), share retesting, maximum scores achieved by retesting, share of students with both parents having any postsecondary education, share of students with high-income parents (above 250K NIS annual household earnings at student ages 14-16), 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 wage (ages 30-32, measured in thousands of 2023 NIS). Samples are defined by age at first test: 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 (580-620).

For these reasons, the analysis focuses on the main sample. Older test-takers are less relevant for the research question because they have limited opportunities to retake the test and meaningfully raise their scores—an essential step for admission to Selective STEM programs. Arab test-takers are underrepresented near the 600 threshold due to large achievement gaps, limiting statistical power for analyzing this group in the regression discontinuity design.

[12] The table also shows that those tested at younger ages typically come from higher socioeconomic backgrounds: 49 percent have parents with postsecondary education and 41 percent have high parental income, compared with 41 and 40 percent, respectively, among older test-takers. They are also more likely to be female—63 percent versus 53 percent among older test-takers.

3 Empirical Strategy

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

Formally, let s_i denote the first score of student i. Let $Z_i = \mathbf{1}\{s_i \geq 600\}$ indicate whether student i scores at or above the round-number threshold, which may cause a perception boost due to left-digit change. For an outcome of interest Y_i, such as applying to Selective STEM degrees, let $Y_i(z)$ denote the potential outcome if student i were assigned to treatment status $z \in \{0, 1\}$. The parameter of interest is the average treatment effect of crossing 600 in the first test, that is: $\tau \equiv \mathbb{E}[\, Y_i(1) - Y_i(0) \,|\, s_i = 600 \,]$.

Identification relies on the assumption that the conditional expectations of the potential outcomes are continuous at the threshold. Formally, for $j \in \{0, 1\}$, $\mathbb{E}[Y_i(j) \mid s_i = s]$ is continuous at $s = 600$. In this context, the assumption requires that no factor determining the outcomes changes discontinuously at the threshold, other than perceived performance. The institutional context supports this assumption: scores are generated through a formal grading process, making manipulation unlikely. Further empirical validation is provided below. Under the continuity assumption, τ is identified and can be expressed as:

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

The discontinuity is estimated using a local linear regression with a 20-point bandwidth ($s_i \in [580, 620]$) and triangular kernel weights centered at 600, which assign greater weight to observations near the cutoff. Because observations at the extreme points receive zero weight, the effective sample is $s_i \in [581, 619]$. Standard errors are heteroskedasticity-robust, and clustered at the score level (Lee and Card, 2008). Robustness checks include bias-corrected estimation (Calonico et al., 2014) and honest inference (Kolesár and Rothe, 2018).[13] 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.

[13]Bias-corrected estimates are based on methods developed for continuous running variables. In this case, the running variable is discrete, but when it has many mass points near the cutoff, local polynomial methods remain appropriate under reasonable assumptions (Cattaneo et al., 2024).

Design validity requires that both the density of test scores and baseline covariates be smooth at 600. Figure G.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 (3 percent of density, $p = 0.06$). However, a parametric check using the (known) approximately normal distribution of scores and fitting a quadratic within a 50-point window finds no significant discontinuity (1 percent, $p = 0.21$). This pattern is consistent with the institutional setting, where manipulation is implausible.

Table G.3 further supports validity by showing balance in baseline covariates using the full test 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 are significant at the 5 percent level, and all discontinuities are small in magnitude. Appendix Table G.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 STEM programs within five years of the baseline test—on all predetermined variables (see Appendix A for more details). Predicted probabilities of applying to Selective STEM, based on all baseline covariates, are smooth at 600 when using all tests and when focusing on the main sample (Figures G.2a and G.2b).[14]

Finally, I confirm that admission to Selective STEM programs does not increase discontinuously at 600. Appendix B provides full details, and Figures B.1 and B.2 show how enrollment in Selective and Lucrative STEM degrees varies with UPET scores. Unlike the main analysis, which uses first-test scores, these figures rely on admission-relevant scores (each individual's highest score) to test whether 600 provides any admission advantage.

Figure B.1 shows that enrollment rates are low and show no discontinuous jump at 600. The slope becomes steeper above 600, indicating that slightly higher scores begin to matter for some, likely those with higher Bagrut GPAs who can qualify with lower UPET scores. When restricting the sample to applicants (Figure B.2), enrollment rates rise substantially, implying that applicants already recognize their stronger admission prospects. Yet again, there is no discontinuous change at 600. This confirms that scoring 600 does not confer a direct admission advantage.

[14] Details on the exact specification appear in Appendix A, and estimated coefficients are reported in Table G.5.

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 test increases applications to Selective STEM degrees at universities. For all students, the regression discontinuity estimate is 1.7 percentage points (an increase from 8.0 to 9.7 percent). Panel (c) focuses on the main sample (age 20 or below), where applications rise by 3.8 percentage points from a 9.6 percent baseline. In contrast, Panel (d) shows no significant discontinuity for older students, consistent with their lower baseline (5 percent) and limited scope to retest before applying.[15]

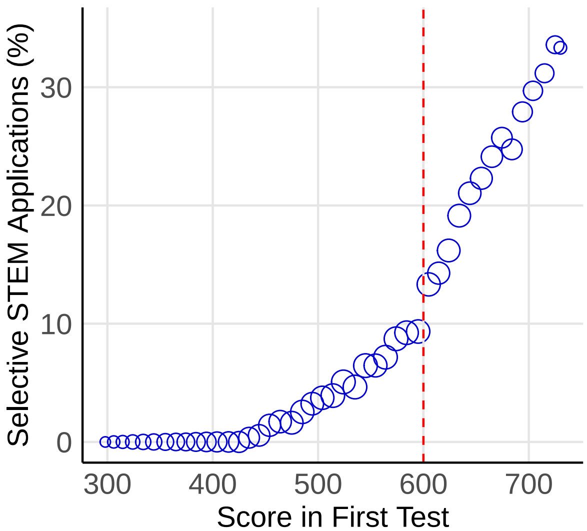
Focusing on the main sample, Columns (2)–(6) of Table 2 show results for additional outcomes. The probability of applying to any Lucrative STEM degree rises by 4.8 percentage points from a baseline of 19.0 percent. This increase is offset by a 2.7 percentage point decline in applications to lower-return 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 56.1 percent—and this estimate is noisy. The average wage associated with chosen programs, at ages 30–32, increases by 740 NIS relative to a baseline of 16,400 NIS (about 200 USD, or 4.5 percent).[16]

Results for other subsamples are reported in Table G.6. As explained earlier, older students have limited opportunities to retest—an essential step for admission to Selective STEM programs. Consistent with this, most estimates show no significant effects. Arab students are underrepresented near the 600 threshold, which limits statistical power. Accordingly, estimates are imprecise but indicate a sizable increase in applications to Selective STEM degrees of about 9.4 percentage points. The baseline rate for this group is already high—23.9 percent—reflecting that Arab students within the sample are top achievers in their communities and therefore more likely to apply to Selective programs. However, this increase comes primarily at the expense of other Lucrative STEM degrees, resulting in no net change in total STEM applications, which already have a high baseline rate of 44.8 percent.

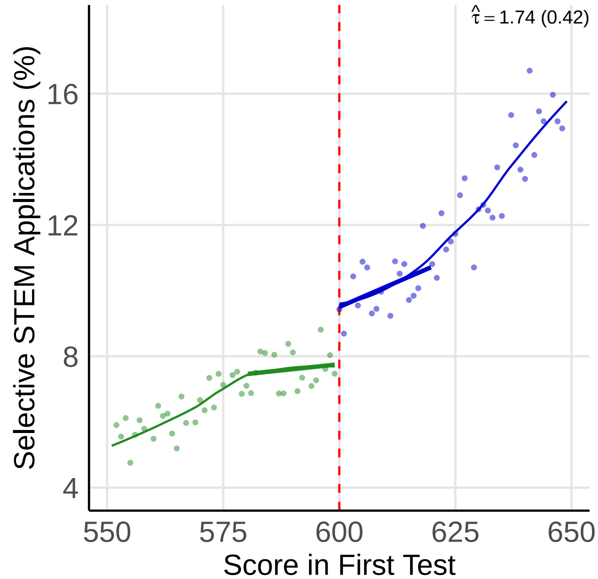
[15]Figures G.2c and G.2d further implement an additional falsification test by estimating the discontinuity at non-round scores. Results show that the discontinuity at 600 is unique.

[16]Additional results for the main sample are presented in Panel A of Table G.7. First, the 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 based on the tech employment rate, the results remain similar. Finally, when extending the application horizon to ever applying, the increase in Selective STEM applications remains similar.

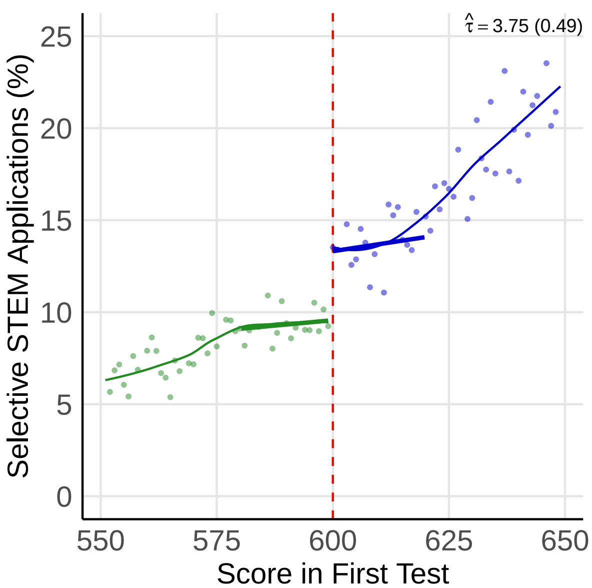
Figure 3: University Entrance Test Scores and Applications to Selective STEM Degrees



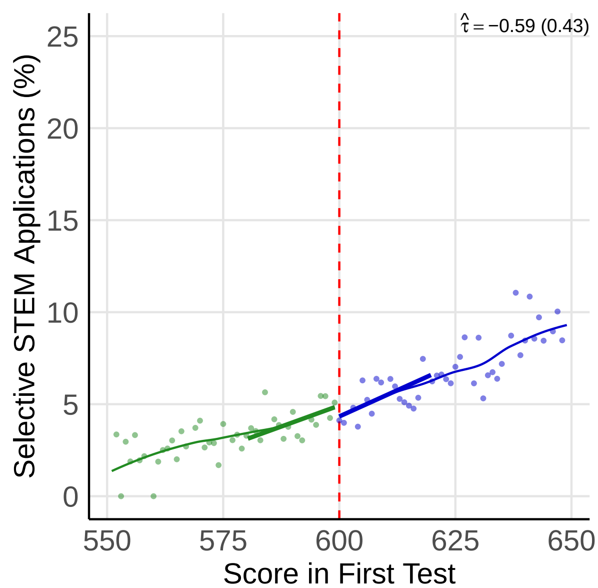
(a) Full Score Distribution



(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 error are 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 threshold. 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 threshold 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. Outcomes in columns (1)-(5) 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 G.1. Other Mid and Other Low classify remaining degrees by associated wage, as described in the text. Any refers to university degrees in any field. Outcome in column (6) 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. The sample size is smaller because associated wage is undefined for individuals who never apply to any program. All indicator outcomes are multiplied by 100. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

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 also validate that the results are robust for other specifications (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 increases applications to STEM much more: 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.[17]

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 parental education. I split the sample into those whose both parents have more than 12 years of schooling, and others. 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 splitting the sample based on parental income (above-median household earnings) or by combining both education and income.

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 with relevant skills to apply, with little effect on others. In this way, the effect reinforces pre-existing differences in STEM readiness.[18]

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 effects are not due to crossing a real 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 4 shows evidence against

[17] Panel B of Table G.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 applying. 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 applying.

[18] 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).

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

	(1)	(2)	(3)
Z: Quantitative Score	Low	High	Diff
Selective STEM	2.31***	8.36***	6.05**
	(0.40)	(2.15)	(2.36)
	[8.16]	[13.75]	
Lucrative STEM	3.00***	10.38***	7.38***
	(0.64)	(2.08)	(2.09)
	[16.73]	[25.89]	
Z: Gender	Males	Females	Diff
Selective STEM	5.15***	2.45***	-2.70
	(1.36)	(0.71)	(1.68)
	[15.54]	[5.83]	
Lucrative STEM	7.80***	2.31*	-5.49**
	(1.66)	(1.18)	(2.20)
	[25.82]	[14.77]	
Z: Parental Postsec Educ	One/None	Both	Diff
Selective STEM	3.10**	4.21***	1.11
	(1.33)	(1.40)	(2.57)
	[8.64]	[10.31]	
Lucrative STEM	2.58	6.38***	3.79
	(1.59)	(1.76)	(3.01)
	[18.03]	[19.84]	

Notes: This table reports the estimated heterogeneity of the effects of crossing the 600-score threshold. Columns (1)–(2) report the estimated impacts (τ from Equation 1) for subgroups with different values of Z, separately. Column (3) reports estimates of the difference in estimated discontinuity effects across subgroups defined by variable Z. Estimation 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 G.1). Heterogeneity dimensions include quantitative advantage (quantitative score at least five points above other domains), female indicator, and high-education 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. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

this. 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 percent (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 4 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.

Together, these results do not support the hypothesis that students view 600 as sufficient for admission. The patterns rule out leading institutional explanations and point to heuristic self-evaluation as the main driver. While institutional factors tied to round numbers cannot be entirely ruled out, there is no evidence of such practices, and it is implausible that a minor or obscure rule could generate effects of this magnitude. Moreover, the evidence indicates that crossing 600 motivates additional effort, consistent with a behavioral rather than institutional response.

The heterogeneity of responses further supports a behavioral interpretation. For students with quantitative strengths, 600 acts as a validating signal that spurs ambition, whereas for others it provides a sense of closure and reduces effort. Such behaviors align with prior evidence on round-number heuristics in test-score responses in other settings (Pope and Simonsohn, 2011; Goodman et al., 2020; Li and Qiu, 2023).

Although the data reveal behavioral patterns, the underlying psychological mechanisms cannot be directly observed without survey information. The response may reflect increased confidence from crossing the threshold or encouragement from peers or family members. To indirectly explore these channels, I examine spillovers to younger siblings' test-taking behavior. Specifically, I test whether the closest-aged sibling (within five years) takes the UPET within three years after the older sibling's first test.

Table 4: Effects of Crossing 600 on Admission-Relevant Outcomes

	Maximum Score	Retest	STEM× Retest	STEM× Max>640	Matriculation Credits	Mat. STEM Credits
	(1)	(2)	(3)	(4)	(5)	(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 threshold 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 581 and 619 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. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Appendix C describes the identification of sibling links and reports the results. Younger siblings are 3.4 percentage points more likely to take the UPET within three years if an older sibling scored just above 600, relative to a baseline rate of 20.9 percent. The effect weakens over longer horizons but remains positive. These findings suggest that not only individuals but also their social circles may exhibit left-digit bias when interpreting test-score signals. This evidence complements prior research 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) and suggests a new channel: siblings' test scores—beyond their educational experiences—can influence younger siblings' decisions.

4.2 Degree Enrollment and Completion

Figure 4 shows that scoring above 600 increases enrollment in a Selective 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 Lucrative STEM program increases by 3.8 percentage points, indicating that applicants who do not enroll in Selective STEM often still shift into other Lucrative STEM degrees.

Additional estimates are shown in Panel A of Table 5. 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 5 reports results for degree completion.[19] Students induced to apply by the perception boost are no more likely to drop out than other students in the sample: Crossing 600 increases Lucrative (Selective) STEM graduation by 3 (1) percentage point from a baseline of 16.7 (6.8) percent. Overall degree completion remains unchanged, with about 86.7 percent graduating. Figure G.3 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.

Additionally, I examine degree achievement using GPA data for graduates in the sample. Two points are worth noting before discussing the results. First, because grading standards differ 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 A). Second, since GPA information is only available for degree completers, this outcome should be interpreted cautiously. To provide context, I also examine the interaction between this indicator (above-median GPA) and a degree in Lucrative STEM, as well as

[19] Table G.8 reports results for the subsamples of older and Arab students, with no significant discontinuities.

results based on pre-degree achievement—the UPET score—to assess degree performance relative to initial rankings.

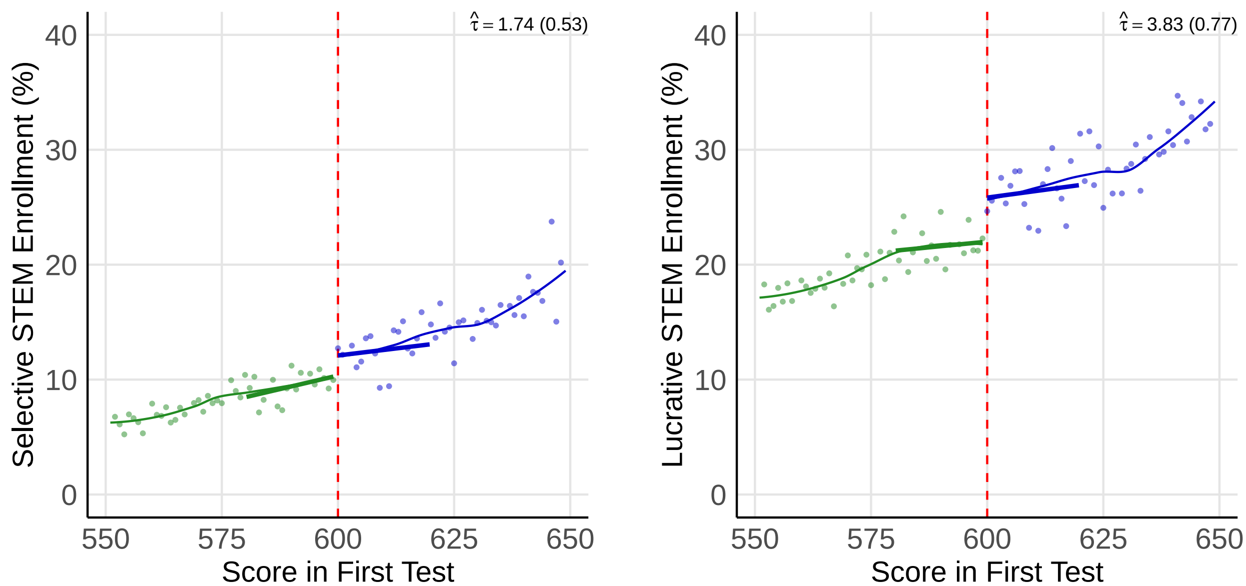
Table 5: Effects of Crossing 600 on Postsecondary Degree Outcomes (Main Sample)

	(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 threshold 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 G.1. Other Mid and Other Low classify remaining degrees by associated wage, as described in the text. Any refers to degrees in any field and institution type (universities or colleges). All indicator outcomes are multiplied by 100. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Table G.9 shows an insignificant 1.7 percentage-point increase in the probability of graduating above the median for those just above 600. When interacting this variable with STEM, there is a 3.4 percentage-point increase from a 10 percent baseline, suggesting that most marginal enrollees do well in their studies. Importantly, when analyzing the pre-degree measure (UPET), no similar pattern emerges, indicating that these students started with relatively lower rankings yet graduated with fair grades. 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 threshold on STEM degree enrollment. The x-axis is the first 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: listed in Table G.1). The sample includes 45,950 first-time test takers (age 20 or younger) in 2000–2009, who 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 error are 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 6. Consistent with the application stage, the effects are concentrated among students with a quantitative advantage: enrollment 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 G.10 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 G.11 and G.12 examine effects by institution type, field, and degree level (undergraduate versus graduate) for both the main sample and the quantitatively advantaged subsample.

Table 6: Heterogeneity of the Effects of Crossing 600 on Degree Enrollment (Main Sample)

	(1)	(2)	(3)
Z: Quantitative Score	Low	High	Diff
Selective STEM	0.65	5.55***	4.90**
	(0.61)	(2.15)	(2.43)
	[9.57]	[13.94]	
Lucrative STEM	1.92*	10.17***	8.24***
	(1.03)	(1.91)	(2.53)
	[19.42]	[29.67]	
Z: Gender	Males	Females	Diff
Selective STEM	2.76*	0.67	-2.09
	(1.49)	(0.65)	(1.87)
	[20.37]	[4.59]	
Lucrative STEM	5.97***	1.62*	-4.35*
	(2.17)	(0.97)	(2.58)
	[37.44]	[12.30]	
Z: Parental Postsec Educ	One/None	Both	Diff
Selective STEM	0.92	2.48***	1.57
	(1.18)	(0.72)	(1.55)
	[9.37]	[11.70]	
Lucrative STEM	3.75***	3.89**	0.14
	(1.29)	(1.57)	(2.41)
	[19.49]	[23.99]	

Notes: This table reports the estimated heterogeneity of the effects of crossing the 600-score threshold. Columns (1)–(2) report the estimated impacts (τ from Equation 1) for subgroups with different values of Z, separately. Column (3) reports estimates of the difference in estimated discontinuity effects across subgroups defined by variable Z. Estimation 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 G.1). 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. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

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 7 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 4 percent.[20] Annual earnings rise by 13,200 NIS, about 6 percent (5 percent).[21]

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 0.25 percent baseline (only marginally significant). Additionally, while overall employment rates (salaried or self-employed) remain unchanged, the likelihood of working in the technology sector increases significantly, by 3.5 percentage points from a baseline of 26 percent.

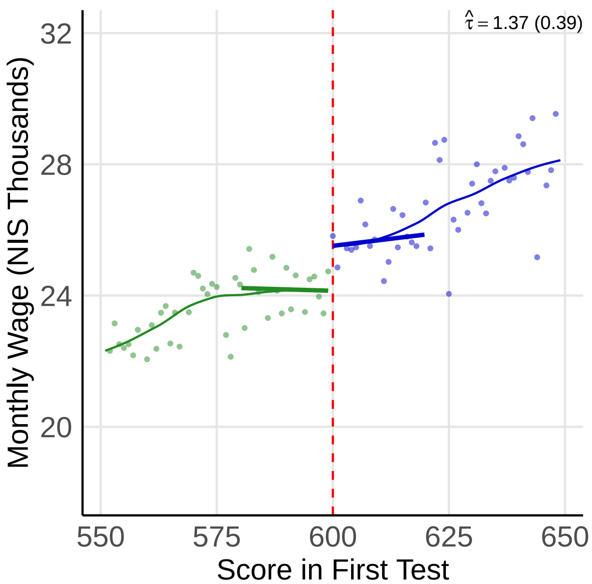
Heterogeneity estimates in Table 8 show that labor-market gains are concentrated among quantitatively advantaged students, consistent with their stronger responses at the application and enrollment stages. Their wages rise by 3,800 NIS, compared to an insignificant 600 NIS gain for others. Similar patterns appear by gender and SES: effects are larger for males and comparable across high- and low-SES students, though these estimates are imprecise. Table G.14 further shows that quantitatively advantaged students experience especially large returns, including higher probabilities of reaching the top of the earnings distribution.

[20] When reporting percentage effects in the main text, I use the log specification because earnings rise significantly at the top of the distribution. Using levels to infer percentage changes would therefore overstate effects at higher incomes. The log specification better captures proportional differences and avoids such exaggeration.

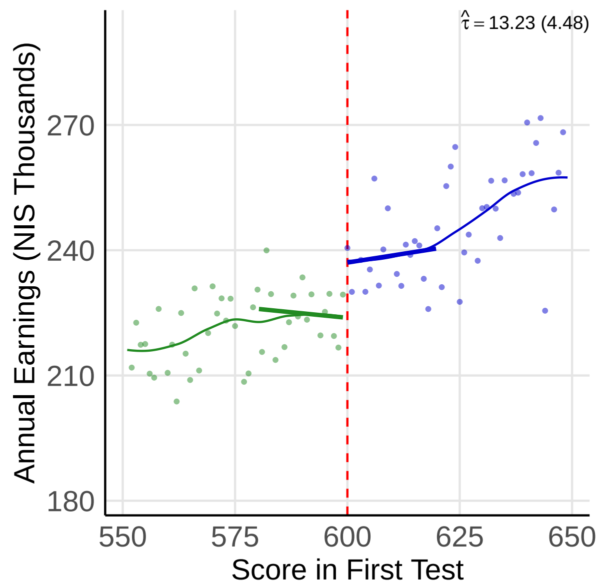
[21] Table G.13 reports results for the two other subgroups. Among Arab students, earnings rise after crossing the threshold, but the increase is statistically insignificant. For older students, by contrast, there is a significant decline in employment rate, likely reflecting their reduced retesting.

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.

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 2022-2023. 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, who 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 error are reported. The estimation sample is restricted to 17,912 test takers who scored between 581 and 619.

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 beginning approximately 10 years 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.[22] 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.

Table 7: 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***	0.05***	1.37***	0.04**
	(4.48)	(0.02)	(0.39)	(0.02)
Mean	223.81	12.24	24.15	9.86
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 threshold 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. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

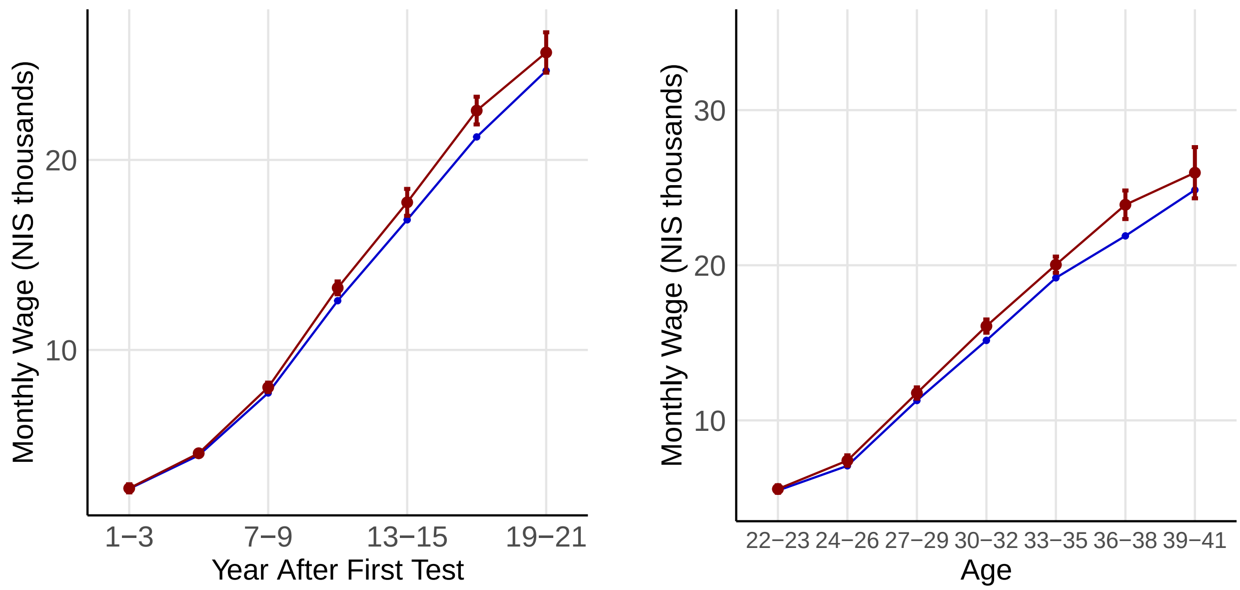
[22] Note that most students work (part- or full-time) during their studies—nearly 60 percent within three years after the test, rising to 76 percent within six years.

Table 8: Heterogeneity of the Effects of Crossing 600 on Labor-Market Outcomes (Main Sample)

	(1)	(2)	(3)
Z: Quantitative Score	Low	High	Diff
Monthly Wage	0.58	3.80***	3.22*
	(0.72)	(1.18)	(1.72)
	[23.20]	[26.84]	
Annual Earnings	5.42	38.67***	33.25**
	(6.15)	(13.37)	(16.11)
	[213.17]	[255.52]	
Z: Gender	Males	Females	Diff
Monthly Wage	1.71***	0.51	-1.20
	(0.59)	(0.41)	(0.81)
	[31.97]	[19.68]	
Annual Earnings	24.08***	2.28	-21.80***
	(7.09)	(3.74)	(7.95)
	[288.30]	[183.44]	
Z: Parental Postsec Educ	One/None	Both	Diff
Monthly Wage	0.91	1.73***	0.82
	(0.65)	(0.56)	(0.93)
	[24.03]	[24.24]	
Annual Earnings	14.65***	12.14**	-2.51
	(5.39)	(5.91)	(7.22)
	[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. Columns (1)–(2) report the estimated impacts (τ from Equation 1) for subgroups with different values of Z, separately. Column (3) reports estimates of the difference in estimated discontinuity effects across subgroups defined by variable Z. Estimation 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 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. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

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 threshold (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. The sample is further restricted to individuals with positive wage income in the relevant period. Individuals are included only if the calendar year in which they reach the relevant age (or years since the test) is before 2023, the last year with available tax data. Later cohorts are excluded from that outcome's analysis.

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 G.4 shows similar patterns for tech employment, suggesting that marginal STEM entrants remain in technology-sector employment 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 G.5 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 G.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. (2014) and Kolesár and Rothe (2018). Results, shown in Table G.15, 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 again drives the results, with similar magnitudes to the main analysis (Table G.16). 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 D 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 of all students 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 thresholds are not analyzed because students at those margins are more likely to apply to colleges, for which we lack complete application data.[23]

[23] 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 college applications, such evidence is only suggestive.

5 Returns to STEM

Having established that crossing 600 increases both STEM enrollment and wages, I now turn to exploiting this behavioral discontinuity in a fuzzy regression discontinuity design to estimate the local returns to STEM for compliers. Recall from Section 3 that s_i denotes the first test score and that crossing the 600 score threshold is defined as $Z_i \equiv \mathbf{1}\{s_i \geq 600\}$, which serves as the instrument in this setting.

Degree choice is a multivalued decision. 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 $D_i^* = d^*$. A complication arises because $D_i^*(z)$ is inherently a multivalued, unordered treatment. Estimating returns separately for each degree program would require multiple instruments (e.g., Kirkeboen et al., 2016).

With a single binary instrument (scoring above 600), I instead binarize the treatment and focus on Lucrative STEM degree enrollment. The next subsection outlines this approach and discusses the conditions under which the binary specification remains valid and provides supporting evidence.

Finally, the last subsection presents results from an alternative fuzzy regression discontinuity approach with a continuous treatment, which assigns each student the expected wage implied by the average wages of their enrolled degrees. This approach relies on a weaker exclusion restriction and captures the average returns to compliers relative to the average observational wage differences between their degree options.

5.1 Identification and Estimation

I focus on the binary treatment of ever enrolling in 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 G.1. 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 threshold:

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

Note that this parameter captures the wage gains from pursuing any Lucrative STEM degree relative to the counterfactual degree choices of compliers. In our setting, these counterfactuals consist mainly of non-STEM degrees in the social sciences and paramedical degrees (see Tables G.11 and G.12), which are associated with substantially lower wages in Israel.

This parameter is identified under the standard assumptions of a fuzzy regression discontinuity design. First, I assume continuity: for each $d \in \{0,1\}$, $\mathbb{E}[Y_i(d) \mid s_i = s]$ is continuous in s at 600, and for each $z \in \{0,1\}$, $\Pr(D_i(z) = 1 \mid s_i = s)$ is continuous at 600. This is supported by the institutional context and by empirical evidence discussed earlier. Second, I assume relevance: crossing the 600 threshold increases Lucrative STEM enrollment. Formally, $\Pr\big(D_i(1) > D_i(0)\big) > 0$. This follows from the sharp discontinuity in STEM enrollment at the threshold documented above. Third, I assume monotonicity, ruling out defiers. Formally, $\Pr\big(D_i(1) < D_i(0)\big) = 0$, that is, 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).[24]

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.

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. This channel is not unique to our setting and could, in principle, arise in contexts with admission-based thresholds. However, the evidence does not support this mechanism: labor-market effects emerge only years later, 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, making the initial score irrelevant in the long run. Finally, subgroups whose degree choices do not change at 600 show no subsequent wage gains.

Second, a potentially more serious violation in our setting is that 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. Results reported throughout the paper suggest that this violation is unlikely to drive the findings. The main change at the threshold

[24] 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 enroll in 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 were unlikely to consider STEM in the first place. 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.

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 threshold with a narrower 10-point bandwidth.

I also verify that the data are consistent with these assumptions. Within a 10-point window around the threshold, I treat the instrument Z_i as if randomly assigned. In this narrow range, the fuzzy regression discontinuity simplifies to a standard instrumental-variable comparison of mean outcomes just above and below the threshold. I apply the joint instrument-validity test of Kitagawa (2015), which uses a Kolmogorov–Smirnov (KS) statistic to test the nonnegativity of compliers' outcome densities. The null of instrument validity cannot be rejected ($p = 0.40$), indicating that the observed data are consistent with the full set of IV assumptions within this specification.

5.2 Main Estimation Results

Table 9 presents the estimation results. I first estimate returns at ages 30–32, the latest point at which all students are observed. The estimated returns are 21,500 NIS in the main sample and 20,000 NIS in the quantitatively advantaged subsample. The latter estimates are preferred because the identifying assumptions are less likely to be violated in this group. The corresponding log estimates, 80 log points, imply roughly a 120 percent wage increase at ages 30–32.

Results are similar, and slightly larger, at older ages (33–35) and in 2022–2023 (ages 30–43). All estimates are statistically significant. The first-stage F-statistics exceed the conventional threshold of 10, supporting the relevance of the instrument. Instrumental-variable estimation based on a kernel-weighted difference in means at the threshold (Table G.17) yields nearly identical results. Using total annual earnings as the outcome (Table G.18) gives comparable magnitudes.

I also follow Abadie (2003) to describe outcomes for the four groups defined by their compliance behavior. Panel B of Table 9, focusing on the quantitatively advantaged subsample, shows that the mean wage of treated compliers at ages 30–32 (30,500 NIS) slightly exceeds that of always-takers (24,300 NIS). Similarly, untreated compliers earn somewhat less (10,500 NIS) than never-takers (14,600 NIS), suggesting that compliers particularly benefit from STEM relative to non-STEM degrees, though these differences are statistically insignificant.

To complement this analysis, I re-estimate the models using tech-sector employment as the outcome. Results in Table G.19 show that most compliers transition from non-tech to tech employment when treated. Among quantitatively advantaged students aged 30–32, tech employment rises from 12 percent for untreated compliers to 89 percent for treated compliers. This pattern explains the somewhat higher wages of compliers relative to always-takers, given that STEM graduates earn more in tech (Table G.2).

Taken together, these results show that the local returns to pursuing STEM in this setting are substantial. They are consistent with earlier evidence that compliers perform at least as well in STEM as other students, in both graduation rates and grades. Their untreated counterparts—students scoring just below 600 who did not apply to STEM—appear constrained by self-perception rather than actual ability, leading to sizable long-term losses in earnings and opportunities.

Unlike admissions-margin studies (e.g., Bleemer and Mehta, 2022), these estimates capture variation in degree enrollment driven by application behavior rather than admission constraints. They highlight the returns for students induced to apply following a boost in self-evaluation and suggest that policies should focus on encouraging uncertain applicants to raise their aspirations and pursue demanding STEM degrees, rather than merely expanding program capacity.

5.3 Robustness to Violations of the Exclusion Restriction

The exclusion restriction in this setting requires that all compliers switch only between STEM and non-STEM programs, with no switching occurring across degrees within either group. Appendix F formalizes this condition.[25] This restriction is non-trivial but can be rationalized here: if ability matters primarily for success in STEM, then an improved self-assessment should mainly shift students into those programs. While not a perfect description of reality,

[25] An alternative sufficient condition is the extensive-margin-only complier assumption discussed in Rose and Shem-Tov (2024).

Table 9: Returns to Lucrative STEM Degrees

	Never Takers	Compliers		Always Takers	
	E[Y(0)]	E[Y(0)]	E[Y(1)]	E[Y(1)]	**LATE**
	(1)	(2)	(3)	(4)	**(5)**
A. Sample: Main					
Ages 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)**
Ages 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)**
Ages 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)**
In 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)**
B. Sample: Quantitatively-Advantaged					
Ages 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)**
Ages 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)**
Ages 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)**
In 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)**

Notes: This table reports estimated returns to 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) at ages 30-32, and log monthly wages at ages 30-32, 33-35, and in 2022-2023. The endogenous variable is an indicator for enrolling in a lucrative STEM degree (fields listed in Appendix Table G.1). 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). $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

it is not overly demanding, since STEM degrees are both more selective, more challenging, and more rewarding than other options available to students in our sample.

To assess the validity of this binarization, I implement a visual diagnostic in the spirit of Andresen and Huber (2021). As detailed in Appendix F, the test exploits a key implication of the binarization conditions. It treats the underlying treatment as continuous—measured by degree-specific average wages—and estimates a series of first-stage regressions for alternative binary definitions, using different thresholds of this continuous measure. The first-stage coefficients should increase up to the chosen threshold and decline thereafter. No strong violations are detected in our setting (Figure G.6).

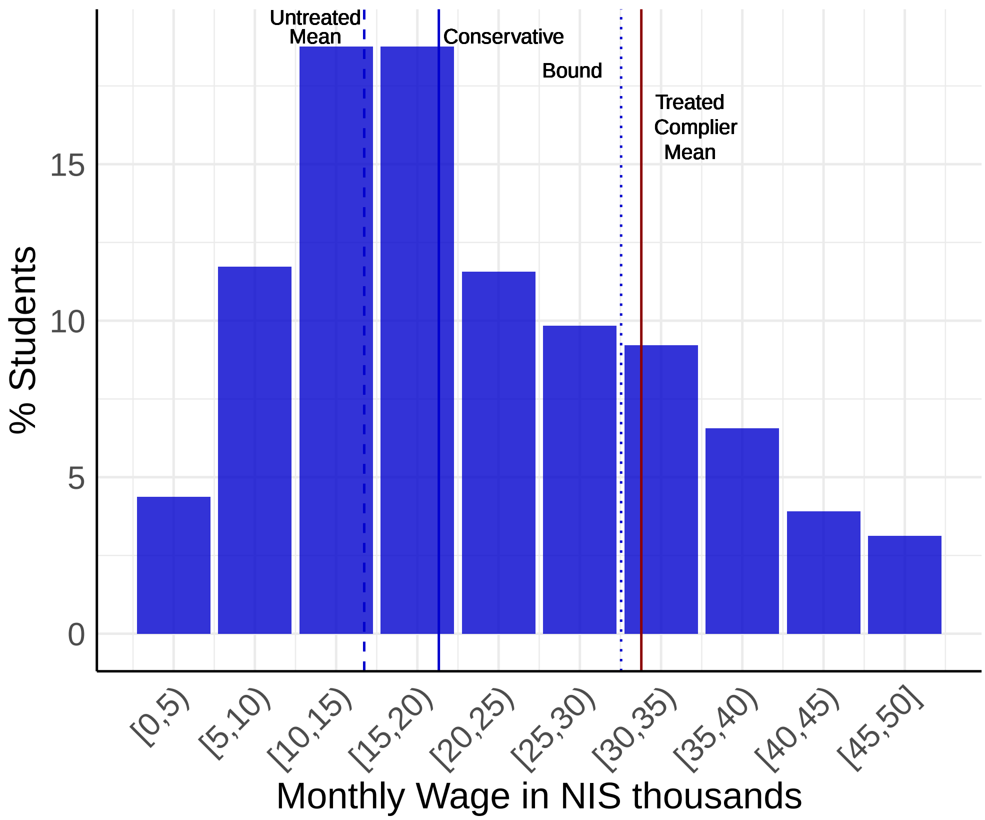
No diagnostic can fully rule out potential violations. To evaluate robustness, I examine the sensitivity of the results to within-group switches that would violate the exclusion restriction. The detailed analysis appears in Appendix F, and the main points are summarized here. First, I argue that switches within the STEM group—cases where $D_i(0) = D_i(1) = 1$ but $D_i^*(0) \neq D_i^*(1)$—are less likely. As shown in Figure G.6, very few students pursue the highest-earning STEM programs, and no compliers appear in that range. Thus, treated compliers are concentrated in a narrow set of STEM degrees, implying that within-STEM switches are unlikely to pose a major concern in our setting.

The more plausible concern is switching within non-STEM degrees. For example, crossing the 600 threshold might shift students from sociology to economics—raising wages without changing the STEM indicator. Such movements violate the exclusion restriction because the instrument affects wages not only through STEM enrollment but also through reallocations across non-STEM programs. The resulting bias would be upward: the reduced form captures wage gains of both binary compliers and within-group movers, while the first stage reflects only the former.

However, following Abadie (2003), I can separately estimate treated and untreated complier means. Even if the exclusion restriction is violated due to within–non-STEM switches, the treated complier mean $\mathbb{E}[Y(1) \mid C]$ remains point identified under the assumption of no within-STEM switches. In contrast, the untreated complier mean $\mathbb{E}[Y(0) \mid C]$ is not point identified in this case and requires additional assumptions or bounds.

To explore this empirically, I construct alternative estimates that relax the exclusion restriction. For simplicity, I focus on a kernel-weighted mean-difference estimation within a 10-point local window around the threshold. Within the subsample of quantitatively advantaged students at ages 30–32, the treated complier mean is 31,300 NIS (Table G.17). 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 threshold, as shown in Figure 7. By monotonicity, students below 600 who do not pursue STEM comprise a mixture of compliers (who would switch at 600) and never-takers. If compliers and never-takers have similar outcomes, the counterfactual mean equals the group average of 14,200 NIS. This implies a LATE of 17,100 NIS—slightly smaller than the 20,000 NIS obtained in the main fuzzy regression discontinuity.

Figure 7: Returns to STEM, Conservative Estimates (Quantitatively Advantaged Subsample)



Notes: The figure shows wage distributions for students not enrolled in any Lucrative STEM degree within the 591–599 window of first test scores (blue). As discussed in Appendix F, 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 (Conservative, dashed blue line), and the mean among those with the highest wages in this group (Bound, 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 students 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 G.1.

Because the shares of compliers and never-takers are point identified from the first stage and from the share of non-STEM students above 600, it is possible to identify the proportion of each group within this sample. A conservative lower bound is obtained by trimming the untreated distribution such that compliers are assumed to be those with the highest wages (following Lee, 2009). This yields worst-case bounds that still imply a positive but small estimate of about 1,200 NIS.

Finally, I construct an alternative conservative counterfactual that trims by degree choice rather than by realized wages. Specifically, I restrict untreated individuals below the threshold to those enrolled in programs with the highest observed mean wages. This specification is conservative, as it compares compliers to students pursuing the most rewarding non-STEM degrees available. The resulting estimates imply sizable gains of about 12,500 NIS.

5.4 Returns Relative to the Average Student

This subsection reports results from a complementary fuzzy regression discontinuity analysis that treats the endogenous treatment as an ordered continuous variable. Recall that degree choice is a multivalued, unordered treatment. In this analysis, I impose an ordering on programs based on their observed average wages, as described in detail in Appendix E. Each program is assigned its mean wage, estimated outside the regression discontinuity sample: $m(d) = \mathbb{E}[Y_i \mid D_i^* = d,\ s_i \notin [581, 619]]$. Each student is then assigned the mean wage of their chosen program, mapping degree choice into an ordered continuous treatment.

This approach offers two key advantages. First, it relies on slightly different assumptions. Second, it provides insight into whether the returns for compliers in this setting are smaller than those of the average student in their chosen degree—a relevant question given that compliers tend to be lower-scoring and initially uncertain about applying.

Beyond continuity, this analysis relies on a positive monotonicity condition, which is the only restriction on how students' degree choices respond to crossing the 600 threshold: the probability of selecting degree programs with higher expected wages should weakly increase for all students, and strictly increase for some. I discuss this assumption and provide supporting evidence in the appendix.

It also requires an exclusion restriction: crossing the 600 threshold affects wages only through changes in degree choice. However, this condition is weaker than the exclusion restriction required earlier, as it allows wages to be influenced by movements between non-STEM degrees as well. This assumption is supported by the evidence in the paper: wage effects emerge only after degree completion and only among students who changed their degree choice.

Under this monotonicity, together with continuity, exclusion, and relevance, the fuzzy regression discontinuity design with m as the endogenous treatment identifies an informative parameter: it compares the wage gains of compliers to the observational wage difference between their chosen and counterfactual programs. As shown formally in the appendix, this estimate likely provides a lower bound for the ratio between complier returns and average student returns, since average wages likely combine causal returns with selection bias.[26]

Appendix E presents the formal derivation and results. The estimates indicate that compliers' gains are not smaller—and likely larger—than the average returns of students in their chosen programs. This finding aligns with the interpretation of earlier results and reinforces the main conclusion: heuristic self-evaluation prevents capable students from pursuing more selective degrees where they could achieve returns at least as large as those of other students.

6 Conclusion

This paper demonstrates that uncertainty about ability and reliance on cognitive heuristics distort the allocation of students into selective and rewarding educational programs within test-based meritocracies. I exploit a regression discontinuity at a round-number threshold in the national university entrance test. The threshold offers no admission advantage, yet crossing it in initial tests sharply raises applications to high-return STEM programs. Because the threshold lies below actual admission cutoffs, the effect reflects left-digit bias—students interpret the round number as a meaningful signal of ability. The perception boost leads them to improve scores through retesting, apply, and pursue high-return STEM degrees, increasing enrollment by about 3.8 percentage points (20 percent).

Tracking outcomes for over two decades reveals large and persistent gains from these heuristic-influenced choices. Monthly wages rise by about 1,400 NIS (4 percent), annual earnings by 13,200 NIS (5 percent), and employment in the technology sector increases by 3.5 percentage points. The rank within the cohort-specific earnings distribution increases by three percentiles, accompanied by higher probabilities of reaching the upper tail of the distribution.

An instrumental-variable analysis exploiting the perception boost as an instrument for STEM enrollment reveals substantial returns, resulting in wage increases of about 120 percent. While these estimates rely on a nontrivial exclusion restriction, supporting evidence suggests

[26]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).

that the data are consistent with it. Even under conservative specifications that relax the exclusion, the implied gains remain large. This policy margin differs from the conventional admission margin, which primarily informs seat expansion policies. The estimates emphasize a complementary margin—motivating capable but uncertain students to apply. However, the results indicate that such interventions are unlikely to close enrollment gaps. The evidence suggests that belief effects arise only where quantitative preparation is already sufficient, implying that reducing gender or socioeconomic disparities likely requires earlier pipeline investments in preparation.

Two considerations affect external validity. First, to achieve clean causal identification in administrative data, the analysis centers on a single round-number threshold and one high-stakes decision. While this isolates a narrow margin, self-evaluation likely affects educational choices more broadly. The estimates should therefore be viewed as a local manifestation of a general behavioral mechanism. Future research could enhance external validity by combining administrative data with surveys or experiments that measure self-beliefs and directly examine the underlying psychological mechanisms. Second, the large STEM returns for lower-testing students are estimated in an economy with strong demand for STEM labor, suggesting that generalization is most relevant for contexts with similarly high demand.

To summarize, this paper demonstrates how uncertainty about one's own ability and cognitive heuristics shape high-stakes educational decisions. The findings expose a key weakness in modern admission systems that rely heavily on test-based meritocracies, where those who aspire to selective programs must compete for higher scores. As a result, uncertainty and heuristic reasoning cause equally capable students to take divergent educational and career paths. These results underscore the need for clearer communication and better guidance to help students interpret test-score signals and make more 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.

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.

Anelli, Massimo, "The returns to elite college education: A quasi-experimental analysis," *Journal of the European Economic Association*, 2020, *18* (6), 2824–2868.

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.

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

Cattaneo, Matias D., Nicolás Idrobo, and Rocío Titiunik, *A Practical Introduction to Regression Discontinuity Designs: Extensions* Cambridge Elements: Quantitative and Computational Methods for Social Science, Cambridge University Press, 2024.

Chetty, Raj, David J. Deming, and John N. Friedman, "Diversifying society's leaders? The determinants and causal effects of admission to highly selective private colleges," Technical Report w31492, National Bureau of Economic Research 2023.

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.

Friedman, John N, Bruce Sacerdote, Douglas O Staiger, and Michele Tine, "Standardized Test Scores and Academic Performance at Ivy Plus Colleges," in "AEA Papers and Proceedings," Vol. 115 American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203 2025, pp. 676–681.

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).

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

Israel Innovation Authority, "The Israel Innovation Report 2021," 2021.

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.

Korvorst, Marjolein and Markus F Damian, "The differential influence of decades and units on multidigit number comparison," *Quarterly Journal of Experimental Psychology*, 2008, *61* (8), 1250–1264.

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.

Mountjoy, Jack, "Marginal returns to public universities," Technical Report, National Bureau of Economic Research 2024.

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, *OECD Economic Surveys: Israel 2023*, Paris: Organisation for Economic Co-operation and Development, 2023. https://www.oecd.org/en/publications/oecd-economic-surveys-israel-2023_901365a6-en.html.

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.

Poltrock, Steven E and David R Schwartz, "Comparative judgments of multidigit numbers.," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 1984, *10* (1), 32.

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.

Rothstein, Jesse M, "College performance predictions and the SAT," *Journal of Econometrics*, 2004, *121* (1-2), 297–317.

Sacerdote, Bruce, Douglas Staiger, and Michele Tine, "How Test Optional Policies in College Admissions Disproportionately Harm High Achieving Applicants from Disadvantaged Backgrounds," NBER Working Paper 33389, National Bureau of Economic Research 2025.

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

Sowell, Thomas, *Affirmative Action Reconsidered: Was It Necessary in the First Place?*, New York: Free Press, 1978.

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.

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

<u>For Online Publication</u>

Appendix A Data Appendix

A.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 sector 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).

A.2 Variable Definitions

University applications. The university application dataset is used to construct indicators for application behavior. These indicators capture applications submitted to different postsecondary degrees 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 sectors. 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.

Predicted likelihood of applying to selective STEM degrees. 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 $s_i \in [581, 619]$. 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.

Appendix B Admission Thresholds

The interpretation of this paper's results rests on the assumption that a score of 600 is not an official admission threshold for selective STEM degrees in Israel. This section provides institutional background and empirical evidence supporting this claim. Official admission thresholds for all degree programs are published online each year, but historical thresholds are not systematically recorded or archived.

I was able to find partial records for specific years and universities. In all available cases, admission thresholds for selective STEM degrees are high, making admission with UPET scores around 600 highly unlikely. Admission is based on a weighted average of the UPET score and the matriculation GPA, both considering only the highest score achieved.

At Tel Aviv University, for example, the admission score is calculated as follows: let M denote the matriculation GPA and U the UPET score. First compute $M' = 10.36M - 410.99$. The final composite score is then given by $A = 0.52(M' + U) - 49.15$, where A represents the university admission score.

Given this formula, for a UPET score of $U = 600$, in 2009–2010, the required matriculation GPAs are $M \approx 113.6$ for computer science ($A = 661$) and $M \approx 111.6$ for electrical engineering ($A = 650$). These values are exceptionally high and rare: in my sample, using a noisy GPA proxy, only about 3% and 8% of individuals reach these levels.[27] Similar requirements appear in other universities and years for both fields.

Some programs may also impose additional subject-specific requirements beyond the composite admission score, such as minimum UPET or matriculation scores in mathematics or other subjects. At Tel Aviv University in 2009–2010, no such additional requirements were listed. At Ben-Gurion University in 2012–2013, all engineering programs required a matriculation certificate with at least 4 credits in mathematics (the minimum is 3) and a UPET score of at least 550. In the Faculty of Natural Sciences, which included computer science at the time, applicants were required to have either 5 credits in mathematics with a score of at least 60 or 4 credits with a score of at least 70, as well as a minimum UPET score of 600. In practice, however, these conditions were not binding for computer science or electrical engineering applicants, since their admission thresholds were substantially higher.

One might wonder whether such official criteria could contribute to the behavioral patterns documented in this paper. This is unlikely. First, these criteria are not salient. Second, they vary across institutions, years, and fields, often without correspondence to round numbers.

[27] Values can exceed 100 because the GPA formula adds 10–30 bonus points for advanced tracks (e.g., five-credit programs).

Third, effects appear in applications to institutions with no such requirements in relevant programs.

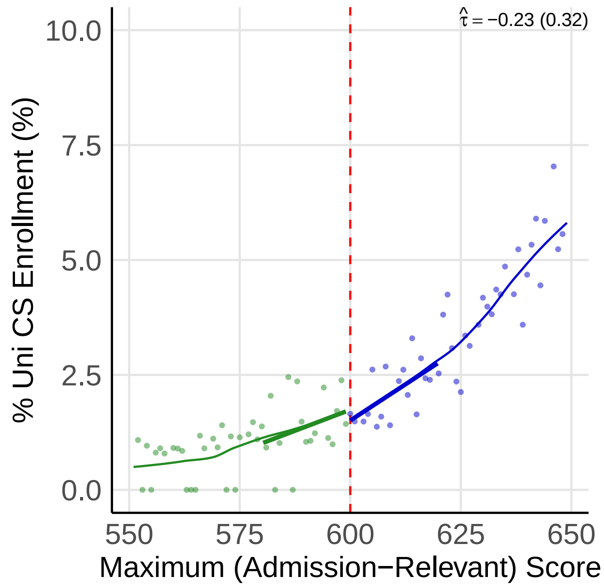
While I do not observe full information for all institutions and years, I use the administrative data to test whether scoring above 600 provides any practical admission advantage. The outcome variable is a dummy for ever enrolling in computer science, electrical engineering, or either of these fields. I estimate the same regression discontinuity as in equation 1, but using each individual's highest test score instead of the first, since admissions are based on the highest score.

Figure B.1 presents the results. Each plot shows enrollment rates by admission-relevant score around 600, along with fitted lines and corresponding regression discontinuity estimates. Panel A focuses on computer science, Panel B on electrical engineering, Panel C on either of the two (selective STEM), and Panel D on all Lucrative STEM degrees. The analysis is limited to university programs due to missing information on applications to colleges. This restriction is minor, since nearly all degrees in this group are university-based. Enrollment rates below 600 are low, and all estimated discontinuity coefficients at 600 are small and statistically insignificant. The slope becomes steeper above 600, indicating that slightly higher scores begin to affect enrollment probabilities—likely among those with higher Bagrut GPAs, who can qualify with lower UPET scores.

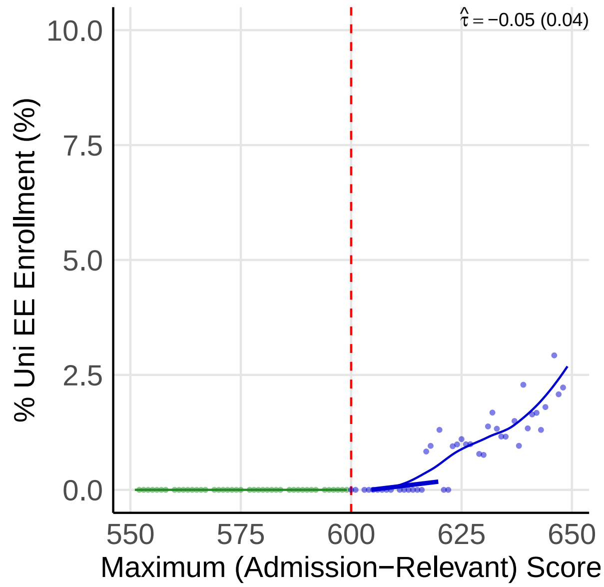
Figure B.2 presents the same analysis, restricted to applicants only. Their baseline enrollment rate is substantially higher, indicating that applicants are already aware of their stronger admission prospects relative to non-applicants. Once again, there is no evidence of a discontinuous jump at 600.

These findings confirm that scoring 600 does not confer a direct admission advantage for selective STEM programs. Moreover, as shown in Section 4, applicants appear aware of this: they frequently retake the test and substantially improve their scores before applying. This supports the interpretation that a formal admissions rule does not drive the results—if such a rule existed, retesting would be unnecessary.

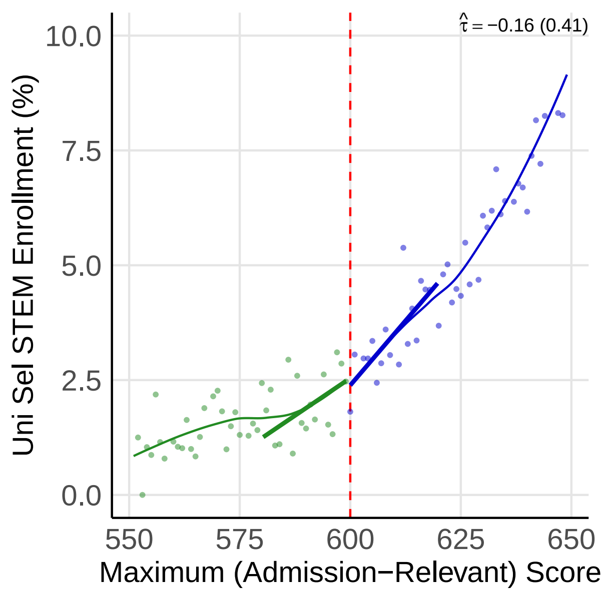
Figure B.1: Selective STEM Enrollment and Admission-Relevant Scores



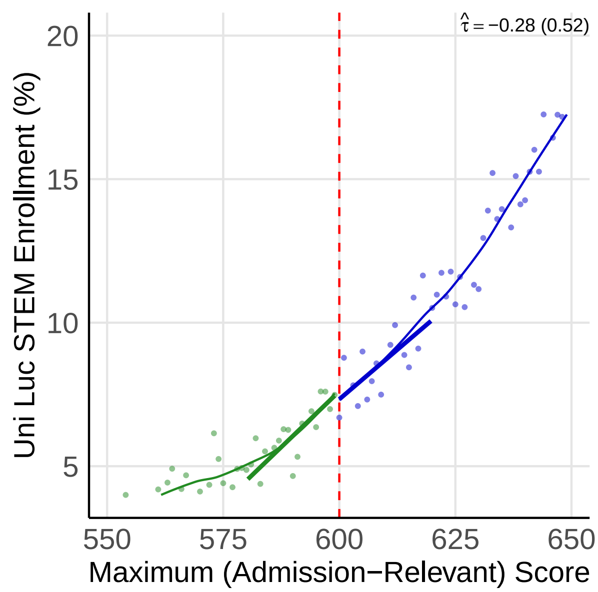
(a) CS



(b) EE



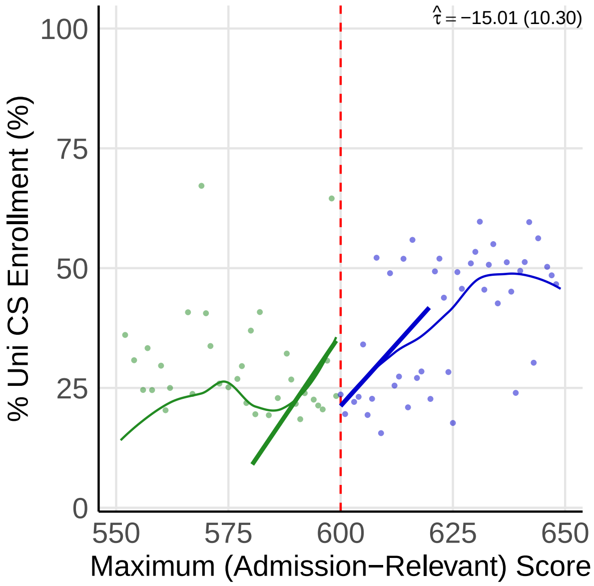
(c) CS or EE



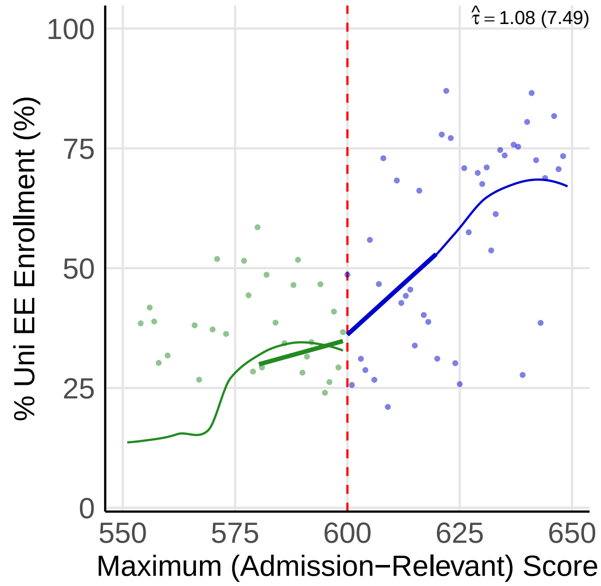
(d) Any Lucrative STEM

Notes: The figures plot the relationship between ever enrolling in a STEM degree (y-axis) and the admission-relevant (highest-ever) university entrance test score (x-axis). The baseline sample includes all individuals in Israel who took their first UPET between 2000 and 2009, restricted to those within 50 points of the 600 threshold ($N = 110,970$). The outcome variables capture enrollment in (a) computer science, (b) electrical engineering, (c) either of these two, or (d) any Lucrative STEM degree listed in Table G.1. All outcomes refer to university programs only. "CS" denotes Computer Science; "EE" denotes Electrical Engineering. All figures show 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 error are reported.

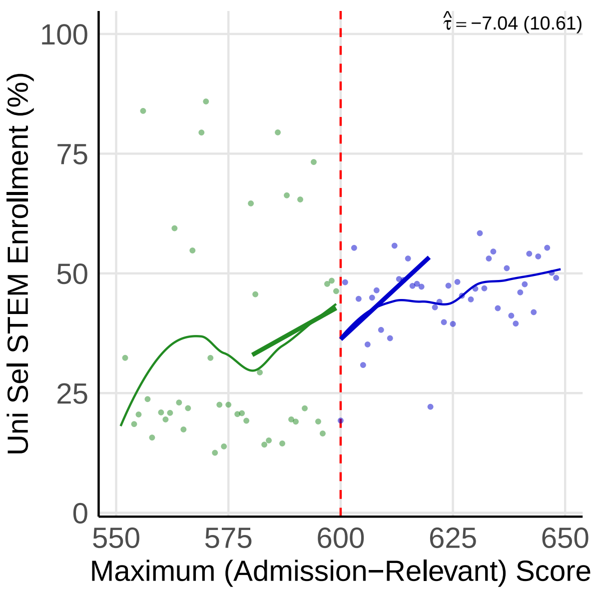
Figure B.2: Selective STEM Enrollment and Admission-Relevant Scores, Conditional on Application



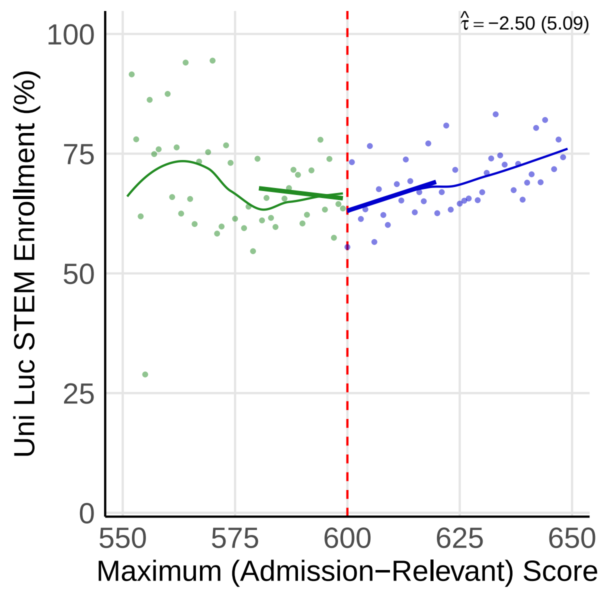
(a) CS



(b) EE



(c) CS or EE



(d) Any Lucrative STEM

Notes: The figures plot the relationship between ever enrolling in a STEM degree, conditional on application (y-axis) and the admission-relevant (highest-ever) university entrance test score (x-axis). The baseline sample includes all individuals in Israel who took their first UPET between 2000 and 2009, restricted to those within 50 points of the 600 threshold ($N = 110,970$). The effective sample in each panel is further restricted to individuals who ever applied to the corresponding program. The outcome variables capture enrollment (conditional on application) in (a) computer science, (b) electrical engineering, (c) either of these two, or (d) any Lucrative STEM degree listed in Table G.1. All outcomes refer to university programs only. "CS" denotes Computer Science; "EE" denotes Electrical Engineering. All figures show 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 error are reported.

Appendix C Young Siblings' Tests

This section examines whether heuristic test-score signals generate spillovers within families. I test whether younger siblings' testing behavior responds to an older sibling scoring just above a round number. Siblings are identified by linking individuals who share the same mother in the population registry. For each student, I retain only one sibling if the age gap is no greater than five years; otherwise, the observation is excluded. For these pairs, I record whether the younger sibling took the test within 3, 5, or 7 years after the older sibling's first test. Using this sample, I estimate Equation 1 with the outcome defined as an indicator for the younger sibling taking the test within three years.

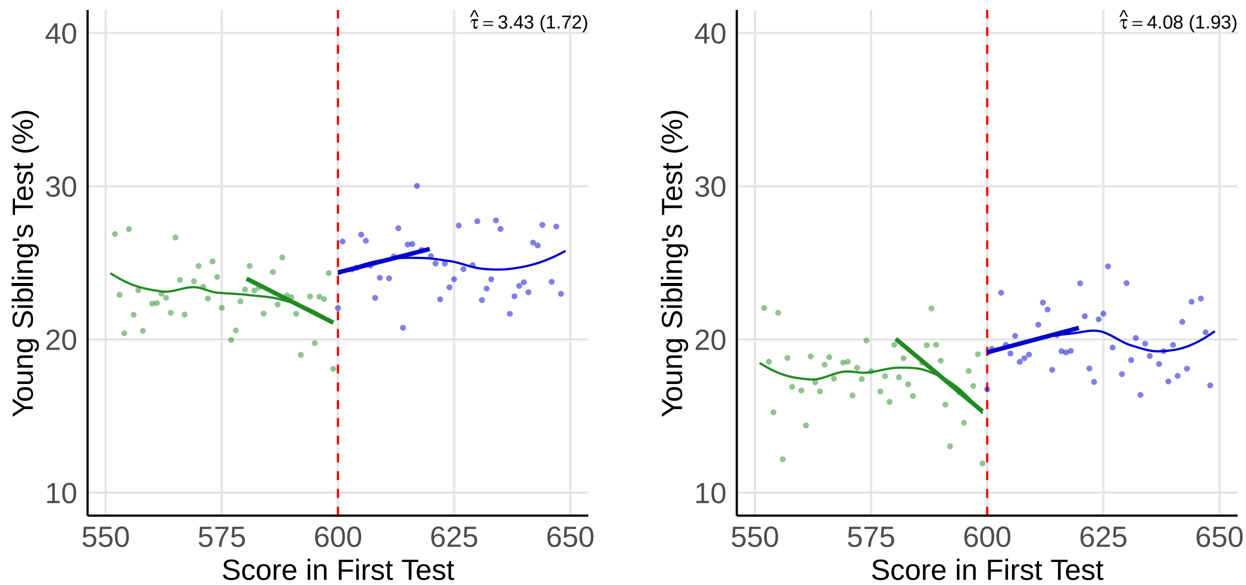
Figure C.1 presents the results. In the full sample (Panel A), younger siblings are 3.4 percentage points more likely to take the UPET within three years if an older sibling scored just above 600, relative to a baseline rate of 20.9 percent. Restricting to the main sample (age 20 or younger) yields a slightly larger estimate of 4.1 percentage points relative to a 15 percent baseline.

Table C.1 reports estimates for additional subsamples and longer horizons. While estimates outside the main sample are positive, they are imprecise and not statistically significant. For example, the estimated effect within the sample of older students is 3 percentage points, which reflects a 13 percent increase from a 23.0 percent baseline. Yet, it is statistically insignificant with a p-value of 0.12.

Moreover, the effect weakens over longer horizons. Using the full test sample, the estimated effect is 2.6 percentage points relative to a 37.7 percent baseline within five years and 1.6 percentage points relative to a 48.3 percent baseline within seven years. While the estimated effects remain positive, they are not statistically significant and may reflect noise.

Overall, these results suggest that heuristic test-score signals produce short-term spillovers within families: younger siblings are more likely to take the test when an older sibling scores just above a round number. Future work could examine the mechanisms driving these spillovers and their persistence over time in settings with larger samples and greater statistical power.

Figure C.1: Effects of Crossing 600 on Young Siblings' Testing



(a) Full Sample (b) Main Sample

Notes: The figure shows the estimated effect of crossing the 600-score threshold on younger siblings' test-taking behavior. The x-axis represents the older sibling's first university entrance test score, and the y-axis indicates whether the younger sibling took the test within three years of that test. Panel A (Panel B) is based on all first-time test takers (the main sample: age 20 or younger) between 2000 and 2009 who scored 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 over the full window. The estimated coefficient τ from Equation 1 and its standard error are reported. The estimation sample is restricted to test takers with scores between 581 and 619. Outcomes are expressed in percentage points.

Table C.1: Effects of Crossing 600 on Young Siblings UPET Testing

	3 Years	5 Years	7 Years
	(1)	(2)	(3)
	A. All		
	3.43**	2.55	1.63
	(1.72)	(2.05)	(1.87)
Mean	20.92	37.68	48.31
N	19,217	19,217	19,217
	B. Main		
	4.08**	2.98	2.31
	(1.93)	(2.35)	(2.48)
Mean	15.12	31.63	48.21
N	8,405	8,405	8,405
	C. Arabs		
	3.21	5.81	2.43
	(7.93)	(6.81)	(6.42)
Mean	45.61	62.06	68.18
N	1,269	1,269	1,269
	D. Older (21 and Above)		
	3.02	1.88	1.02
	(1.91)	(2.77)	(2.28)
Mean	22.97	40.02	46.07
N	9,543	9,543	9,543

Notes: This table reports the estimated effect of crossing the 600 score threshold 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 within three, five and seven years. Dummies are multiplied by 100. The sample includes all first-time UPET takers in Israel from 2000 to 2009 who scored between 581 and 619 on their first test with any young sibling. Panel A reports results for the full sample and Panels B–D present results for subsamples: the main sample (age at test 20 or below), older and Arab test-takers. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Appendix D 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 students who ever take the UPET 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 D.1 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 threshold 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.

Table D.1: 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 threshold. 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). $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Appendix E Returns Relative to Average

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 $D_i^* = d$.

First, I assume continuity: for all programs $d \in \{d_1, \ldots, d_K\}$, the conditional expectation $\mathbb{E}[Y_i(d) \mid s_i = s]$ is continuous in s at 600. Second, I assume exclusion restriction: crossing the 600 threshold affects outcomes only through degree choice, such that $Y_i(d, z) = Y_i(d)$ for all d and $z \in \{0, 1\}$.

Then, I 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 degrees 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, Figure G.6 plots the marginal distributions of degree-associated wages above and below the 600 threshold, estimated via regression discontinuity. That is the estimated share of students above various wage thresholds (equivalent to first-stage estimates under different binarizations). Estimates indicate 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 threshold, 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 threshold. 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 compliers.

Moreover, to probe sensitivity to such a scenario, 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 categories.

Table E.1 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 and potential violations of monotonicity. Thus, I focus only on the subsample of quantitatively advantaged students. The estimates 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 enrolled in 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.

Table E.1: Returns to Degrees Relative to Average Student

	Estimate
Raw Levels	1.64***
N=3,900, F=17.5	(0.47)
Raw Log	0.68**
N=3,900, F=10.8	(0.30)
Controls Log	1.91***
N=3,900, F=12.4	(0.59)
Controls Levels	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 regression that adjusts for predetermined covariates (denoted Controls). Outcomes are monthly wages at ages 30–32 (the same period in which m is measured). Both outcomes and treatments are measured similarly, in log or level as explicitly indicated. The sample is based on 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. The estimation is restricted to quantitatively advantaged students to reduce concerns about the validity of the underlying assumptions (see Section 5). $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Appendix F 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 E. 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. In the main text, it is defined as $D_i = \mathbf{1}\{D_i^* \in D^s\}$, where D^s denotes the set of degrees listed in Appendix G.1. It is equivalent to the following definition, 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 threshold values, the estimated first stage should increase up to the chosen threshold (17.5 in our case) and then decrease. Panels (a) and (c) of Figure G.6 plots the marginal distribution of degree-associated wages above and below the 600 threshold, based on a series of regression discontinuity estimations, and shows evidence that this holds in our setting. The figure shows a monotonic increase in the share of students pursuing degrees above the threshold, with a 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.

Alternative estimates. To construct alternative estimates, I first note that the assumption $\Pr(C_s) = 0$ is relatively weaker in this setting, since the Lucrative STEM category includes only a few highly rewarding and similar degrees. It seems less likely that recieving a perception boost would move students between different engineering programs. Moreover, all programs in this group are higlhy rewarding and offer opportunities in the tech industry. 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 may increase their wages but leaves the STEM indicator unchanged. 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.

Thus, we maintain $\Pr(C_s) = 0$ and allow $\Pr(C_n) > 0$. This allows identifying the treated complier mean using the modified outcome $\widetilde{Y}_i = Y_i D_i$:

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

Never-STEM compliers do not contribute because $D = 0$ for them on both sides, hence the fuzzy regression discontinuity with the 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 $\mathbb{E}[Y(D(1)) \mid C]$. For simplicity, I estimate this and the following parameters using kernel-weighted averages within a narrow 10-point window around the 600 threshold. As shown earlier, this approach yields results nearly identical to the main local linear regression discontinuity specification. Within the quantitatively advantaged sample, the estimate for the treated complier mean is 31,300 NIS.

Our target is the LATE parameter, $\mathbb{E}[Y(D(1)) - Y(D(0)) \mid C]$. Given $\mathbb{E}[Y(D(1)) \mid C]$, it suffices to recover $\mathbb{E}[Y(D(0)) \mid C]$. We approximate this using the outcome distribution of untreated individuals just below the threshold. Under the fuzzy regression discontinuity assumptions, students below 600 who do not pursue STEM consist of a mixture of compliers—who would switch at 600—and never-takers. Figure 7 displays their wage distribution for the quantitatively advantaged subsample. If the two groups have similar outcomes and no systematic differences, the group average provides a consistent estimate of the untreated complier mean. This estimate is 14,000 NIS, implying a LATE of 17,000 NIS—qualitatively similar to the main estimate of 20,000 NIS.

However, we can also identify their shares. 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 with respect to STEM. Let $p_C = \Pr(C)$ which is identified by the first stage. The complier share among units with $D_i = 0$ just below the threshold is then $\omega_C = \frac{p_C}{p_{0-}}$.

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 (as in 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. Using this method, I estimate worst-case bounds of 1,200 NIS in the quantitative advantaged subsample.

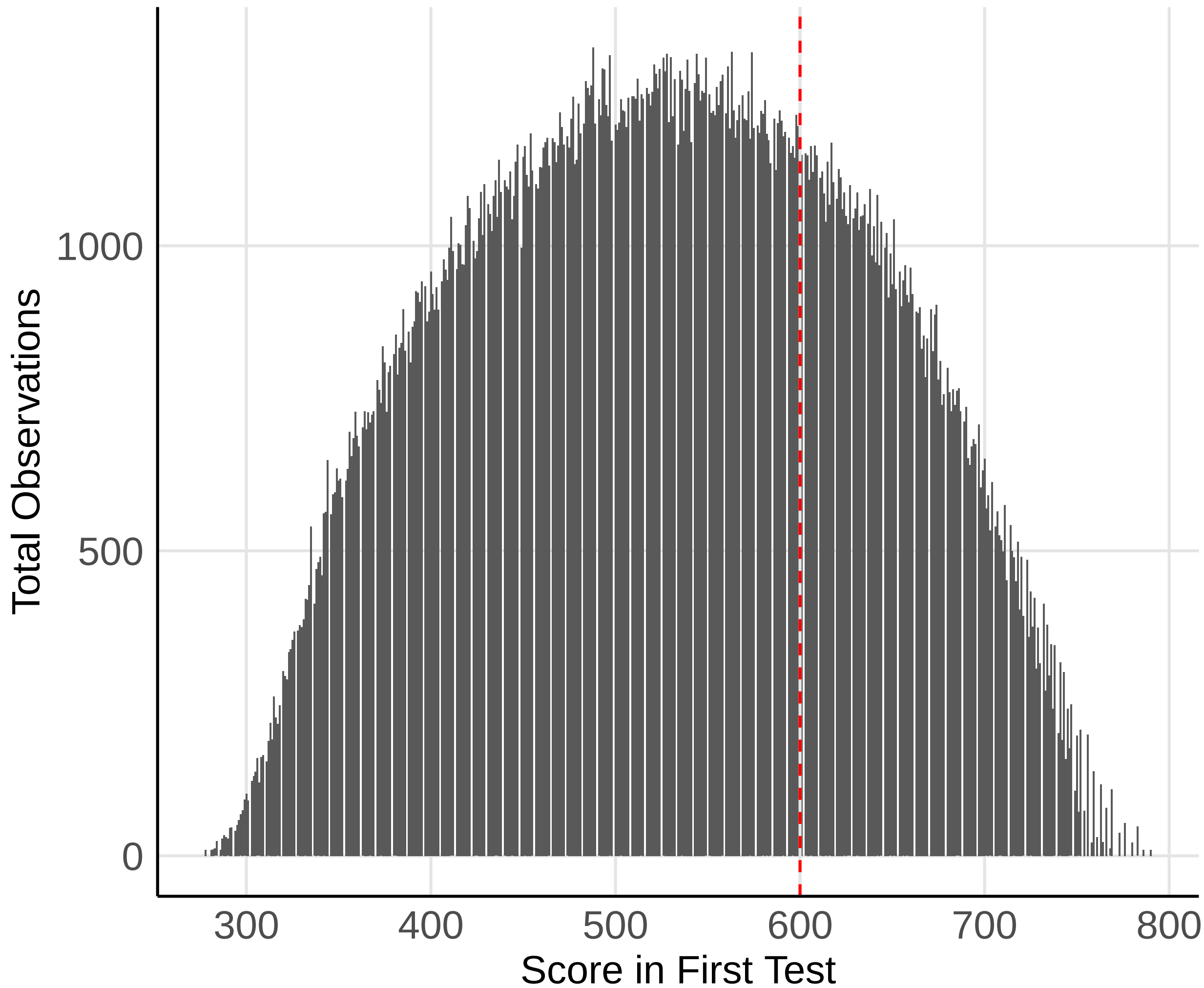
I next propose a tighter yet conservative trimming approach, which relies on the program index m rather than actual wages. Define $Q^-_{m|D=0}(q)$ as the left-limit quantile of m among untreated units ($D = 0$). Under the worst-case ordering, where all compliers lie strictly above never-takers in the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

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

This approach does not provide an upper bound for the untreated complier mean in all cases. For example, compliers could have lower counterfactual outcomes but be especially talented and therefore earn more than the average graduate of their degree programs. Nevertheless, this remains a conservative estimate: within the mixture of compliers and never-takers, it assumes that compliers are those who would have enrolled in the most rewarding degrees. Using this trimming-based approximation, the estimated counterfactual is 19,800 NIS, implying a LATE of 12,500 NIS.

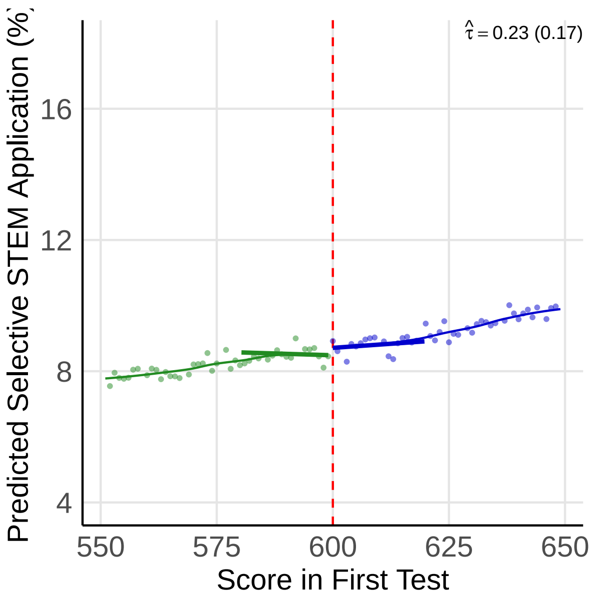
Appendix G Additional Figures and Tables

Figure G.1: University Entrance Test Score Distribution, First Tests

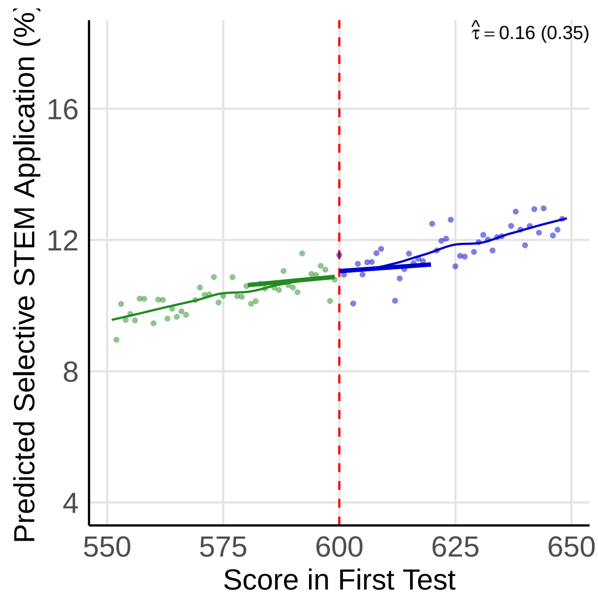


Notes: The figure displays the distribution of first-test university entrance test scores for 359,122 individuals who took their first test between 2000 and 2009. The red dashed line indicates the 600 score threshold. A local linear density test (McCrary, 2008) detects a small, marginally significant 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.

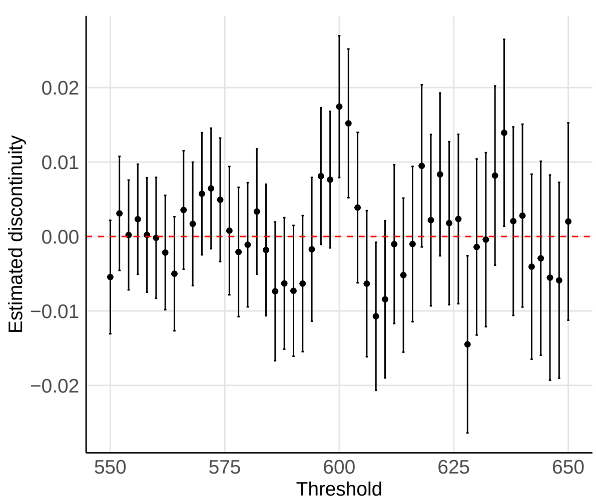
Figure G.2: Regression Discontinuity Design Falsification Tests



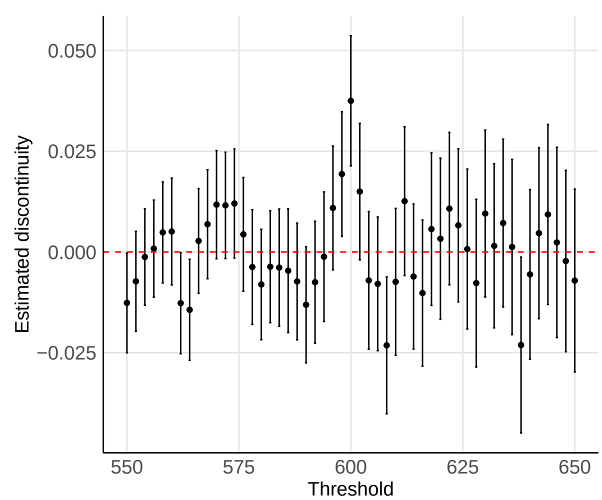
(a) Predicted outcome, all students



(b) Predicted outcome, main sample



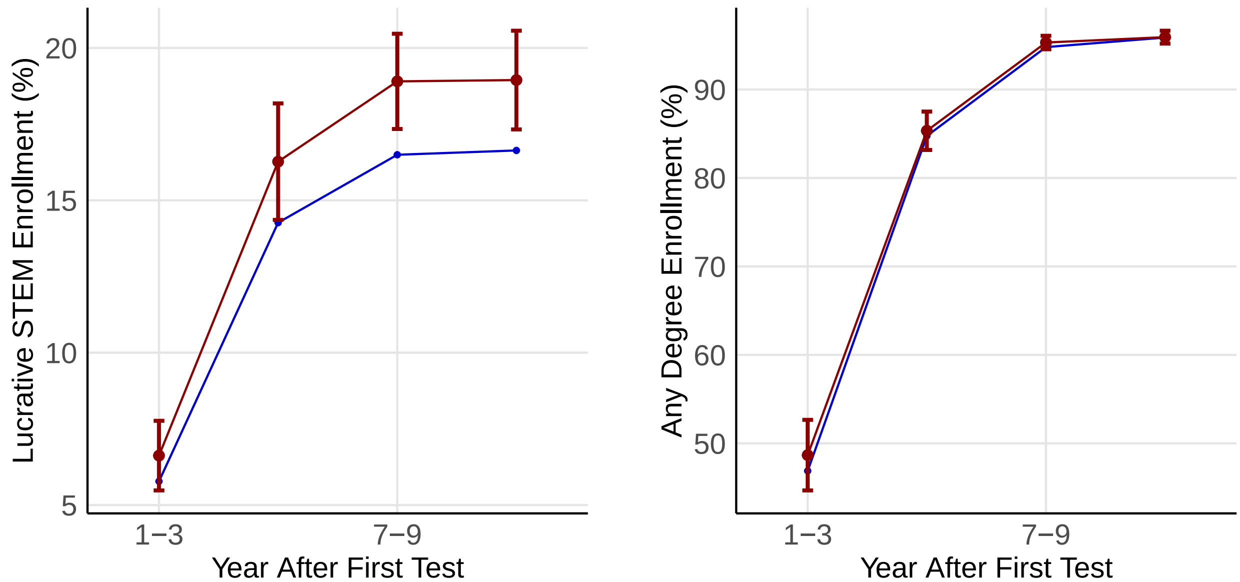
(c) Non-round thresholds, all students



(d) Non-round thresholds, main sample

Notes: These panels present falsification tests. Panels (a)–(b) show estimated discontinuities at the 600 threshold 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 thresholds, 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 581–619 (for panels a–b) or within 20 points of the tested threshold (for panels c–d). Standard errors are clustered by score.

Figure G.3: Degree Enrollment Effects of Crossing 600 Over Time (Main Sample)



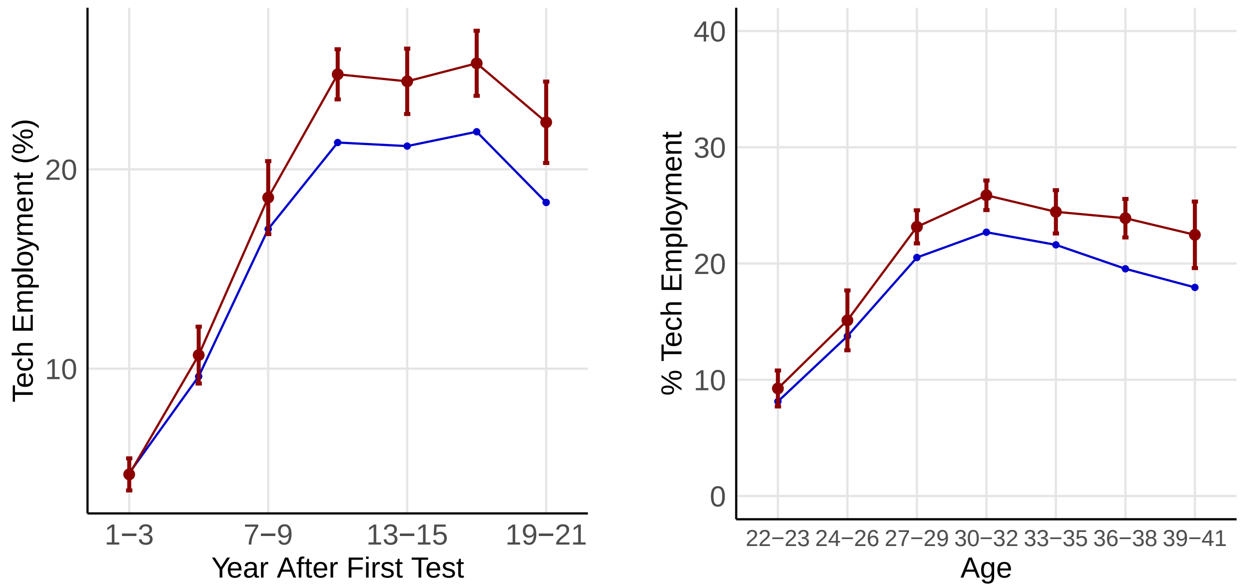
(a) STEM (%)

(b) Any (%)

Above 600 Below

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 G.4: Effects of Crossing 600 on Tech Employment Over Time (Main Sample)



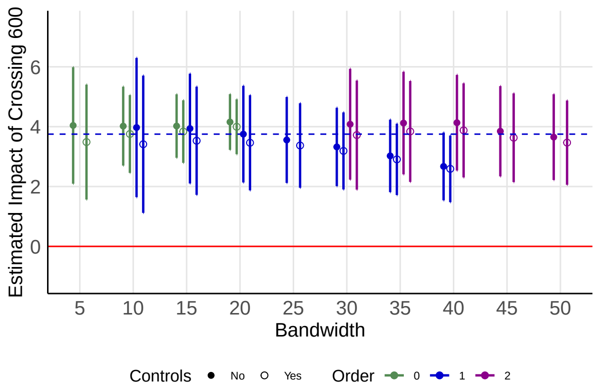
(a) Period relative to test

(b) Age

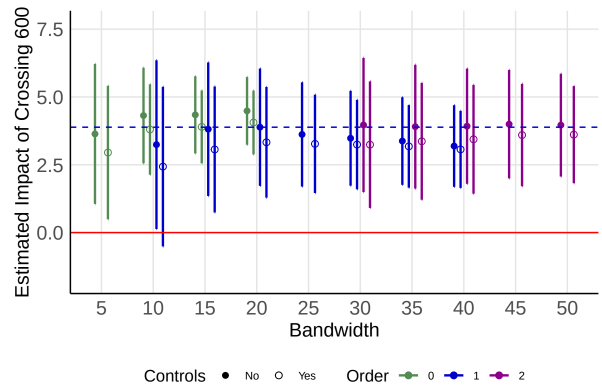
Above 600 Below

Notes: This figure shows baseline means (blue solid line) and implied means above the 600 threshold (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. Individuals are included only if the calendar year in which they reach the relevant age (or years since the test) is before 2023, the last year with available tax data. Later cohorts are excluded from that outcome's analysis.

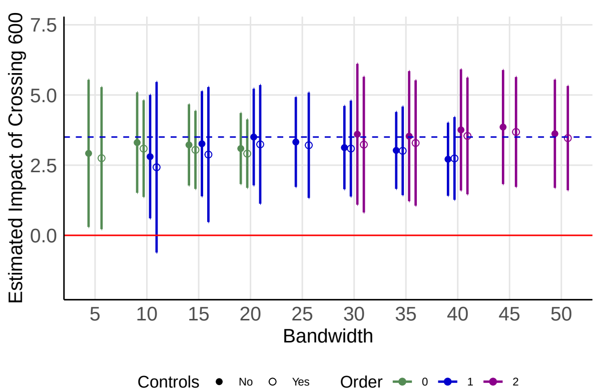
Figure G.5: Robustness of Regression Discontinuity Estimates (Main Sample)



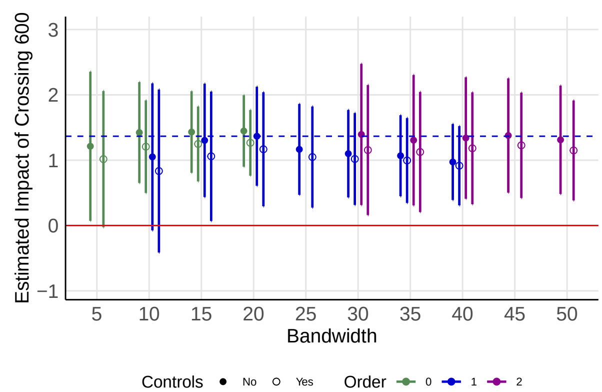
(a) Selective STEM Application (%)



(b) Lucrative STEM Enrollment (%)



(c) Tech Employment (%)

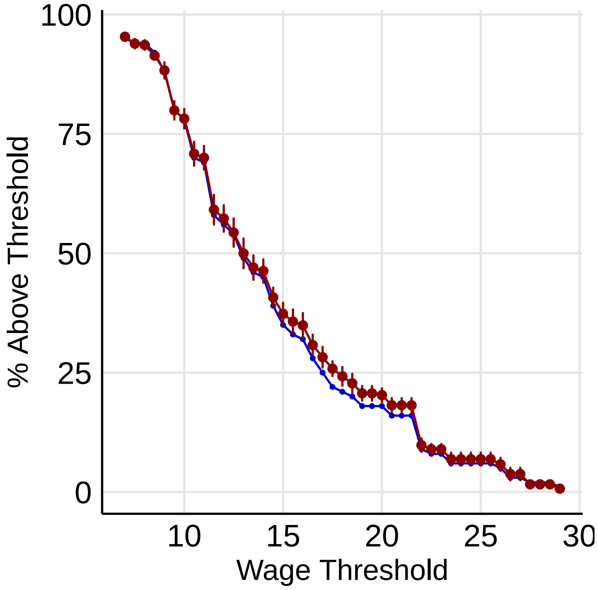


(d) Monthly Wage (NIS thousands)

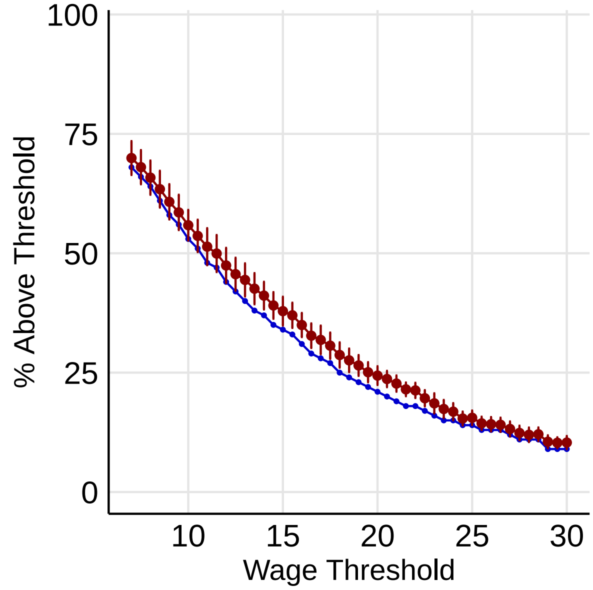
Notes: Each panel shows estimates of the impact of crossing the 600 score threshold under alternative specifications. Points are regression discontinuity coefficients (τ from Equation 1) estimated using different bandwidths (x-axis), polynomial orders, and sets of control variables (detailed in the text). The sample includes 17,912 first-time test-takers in 2000–2009, age $\leq$20 at first test. 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 G.6: Effects of Crossing 600 on Degree-Linked and Individual Wage Thresholds

Main Sample

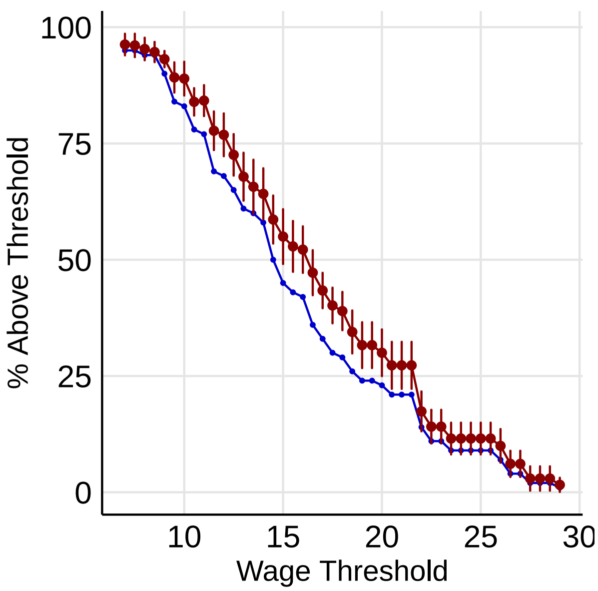


(a) Degree

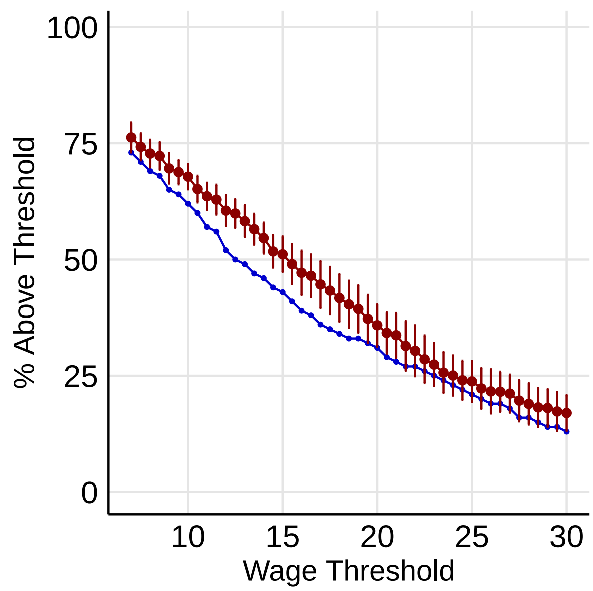


(b) Actual

Quantitatively Advantaged Sample



(c) Degree



(d) Actual

Above 600 Below

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. Panels (a) and (c): outcome is an indicator for selecting a degree with associated average earnings above the threshold on the x-axis. Panels (b) and (d): outcome is an indicator for the individual's own wage exceeding the threshold. Effects are measured in percentage points (y-axis). The sample in panels (a) and (b) includes first-time UPET takers in 2000–2009, age 20 or younger, with first-test scores between 581 and 619. The samples in panels (c) and (d) are restricted to the subsample of quantitatively advantaged students. Standard errors are robust and clustered by score.

Table G.1: Lucrative STEM Degrees

Field	Type (1)	Students (2)	% Tech (3)	Wage (4)
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 wage 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 581-619) to avoid mechanical overlap with the treatment sample.

Table G.2: Average Earnings by Postsecondary Degree and Sector

	N	Sector		Earnings (k NIS)			
		Sector	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
Some College	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 sector. 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 (at ages 30-32) 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 G.3: Falsification Tests: "Effects" of Crossing 600 on Baseline 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	Verbal
	-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 threshold 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 G.4: Falsification Tests: "Effects" of Crossing 600 on Baseline 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	Verbal
	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 threshold 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 G.5: Predictors for Selective STEM Applications

Variable	Estimate	(Std. Error)
Intercept	-2.44***	(0.24)
Arab	-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 581-619) are excluded to avoid mechanical overlap with the treatment sample. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

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

	Selective STEM (1)	Lucrative STEM (2)	Other Mid (3)	Other Low (4)	Any (5)	Associated Wage (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 G.1); non-STEM degrees classified by associated wage; 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 G.7: Effects of Crossing 600 on University Applications, Alternative Outcomes

	CS (1)	EE (2)	Tech (3)	STEM Ever (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 threshold 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 581 and 619 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 G.8: 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 threshold 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 percentage points. The baseline sample includes all first-time UPET takers in Israel from 2000 to 2009 who scored between 581 and 619 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 G.9: Effects of Crossing 600 on Degree Achievement

	Degree GPA		UPET	
	Above Median	Above Median ×STEM	Above Median	Above Median ×STEM
	(1)	(2)	(3)	(4)
	A. Main			
	1.70	3.36***	-2.46	0.60
	(1.88)	(0.77)	(1.67)	(0.55)
Mean	52.62	9.96	29.49	1.49
N	12,733	12,733	12,733	12,733
	B. Main, Quantitative Advantage			
	0.10	7.28***	-6.69**	0.88
	(3.71)	(2.41)	(2.96)	(1.12)
Mean	55.64	15.11	24.02	1.88
N	3,489	3,489	3,489	3,489
	C. Main, No Quantitative Advantage			
	2.27	2.02**	-1.04	0.51
	(2.17)	(0.79)	(1.96)	(0.59)
Mean	51.56	8.16	31.42	1.35
N	9,244	9,244	9,244	9,244

Notes: This table reports the estimated effect of crossing the 600 score threshold 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 in columns (1)–(2) are an indicator for being above the median GPA rank among degree graduates and its interaction with a degree in Lucrative STEM. Outcomes in columns (3)–(4) are an indicator for being above the median UPET rank among degree graduates and its interaction with a degree in Lucrative STEM. All outcomes are defined at the field–institution–cohort level. Indicator outcomes are multiplied by 100, such that coefficients represent percentage points. The baseline sample includes all first-time UPET takers in Israel from 2000 to 2009 who scored between 581 and 619 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 G.10: Effects of Crossing 600 on Postsecondary Degree Outcomes (with Quantitative Advantage)

	(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 threshold 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 wage, as described in the text. Any refers to degrees in any field and institution type (universities or colleges). All indicator outcomes are multiplied by 100. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Table G.11: 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 threshold 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 G.12: Effects of Crossing 600 on Degree Enrollment, Additional Outcomes (with Quantitative Advantage)

	(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 threshold 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 G.14: Effects of Crossing 600 on Labor-Market Outcomes in 2022-2023 (with Quantitative Advantage)

	(1)	(2)	(3)	(4)
A. Earnings (NIS Thousands)				
	Annual Earnings	Log Annual	Monthly Wage	Log Monthly
	38.67***	0.12**	3.80***	0.13**
	(13.59)	(0.06)	(1.20)	(0.05)
Mean	255.52	12.39	26.84	9.99
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 threshold 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 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 G.13: 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 threshold 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 581 and 619 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 G.15: 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 threshold using alternative regression discontinuity methods. Panel A presents robust bias-corrected estimates following Calonico et al. (2014), 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 G.1), 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 G.16: 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)			
	-1	1	5
	(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 G.1), tech employment, and monthly wage (in thousands of 2023 NIS). $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Table G.17: Returns to Lucrative STEM Degrees, Kernel-Weighted Difference

	Never Takers	Compliers		Always Takers	
	E[Y(0)]	E[Y(0)]	E[Y(1)]	E[Y(1)]	**LATE**
	(1)	(2)	(3)	(4)	**(5)**
A. Sample: Main					
Ages 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)**
Ages 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)**
Ages 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)**
In 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)**
B. Sample: Quantitatively-Advantaged					
Ages 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)**
Ages 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)**
Ages 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)**
In 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)**

Notes: This table reports estimated returns to lucrative STEM degrees (δ from Equation 2, estimated as a kernel-weighted difference in means at the threshold). 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) at ages 30-32, and log monthly wages at ages 30-32, 33-35, and in 2022-2023. The endogenous variable is an indicator for enrolling in a lucrative STEM degree (fields listed in Appendix Table G.1). 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 G.18: Returns to Lucrative STEM Degrees in Annual Earnings

	Never Takers	Compliers		Always Takers	
	E[Y(0)]	E[Y(0)]	E[Y(1)]	E[Y(1)]	**LATE**
	(1)	(2)	(3)	(4)	**(5)**
A. Sample: Main					
Ages 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)**
Ages 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)**
Ages 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)**
In 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)**
B. Sample: Quantitatively-Advantaged					
Ages 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)**
Ages 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)**
Ages 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)**
In 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)**

Notes: This table reports estimated returns to 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) at ages 30-32, and log monthly wages at ages 30-32, 33-35, and in 2022-2023. The endogenous variable is an indicator for enrolling in a lucrative STEM degree (fields listed in Appendix Table G.1). 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 G.19: Local Effect of Lucrative STEM Degree Enrollment on Tech Employment

	Never Takers	Compliers		Always Takers	
	E[Y(0)]	E[Y(0)]	E[Y(1)]	E[Y(1)]	**LATE**
	(1)	(2)	(3)	(4)	**(5)**
A. Sample: Main					
Ages 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)**
Ages 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)**
In 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					
Ages 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)**
Ages 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)**
In 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 sector, 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 G.1). 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). $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$