The Importance of Round Test Scores: Human Capital Decisions and Labor Market Outcomes

Preliminary and Incomplete

Yoav Goldstein*

19 September 2022

Abstract. I examine how high-stakes test scores affect university field of study decisions in Israel. I implement a regression discontinuity design around a round score cutoff, 600, which is usually just below the admission threshold for computer sciences, one of Israel's most rewarding university programs. In the short term, I find a 35% discontinuous increase in the rate of applications to the program at 600. Moreover, I find that individuals who score just above 600 retake the test and improve their test scores as well as improve their high-school outcomes before applying to the program, which are both ways to improve admission chances. In the long term, individuals who score just above 600 attain more computer sciences degrees and earn significantly more. The results highlight that an uninformative signal might have far-reaching consequences on human capital investment decisions and labor market outcomes.

Keywords: Educational choice, major, high stakes tests, left-digit bias
JEL Codes: I2, I23, D83, J24

*The Eitan Berglas School of Economics, Tel Aviv University, yoavg2.mail.tau.ac.il. I thank the Central Bureau of Statistics for providing access to data I use in this study at its protected research room in Tel Aviv. I am grateful to my supervisors, Analia Schlosser and Itay Saporta-Eksten for their guidance and support. I also thank Oren Danieli, Dor Leventer, Kobi Mizrahi, Shahar Lahad, Matan Kolerman, Ity Shurtz, Dan Zeltzer, Muly San, and participants in the Israeli Economic Association Annual Conference, HCEO–IESR Virtual Summer School On Socioeconomic Inequality, Israeli National Institute for Testing and Evaluation Seminar, and HCEO-FAIR Summer School On Socioeconomic Inequality for useful comments and suggestions.

1 Introduction

Test scores have significant consequences on human capital and labor market outcomes. One important channel for how test scores affect outcomes is by providing ability signals (e.g., Stinebrickner and Stinebrickner, 2012; Avery et al., 2017). The prior literature focused on rational responses to test scores without paying attention to the role of heuristics. However, the well-documented left-digit bias implies that the leftmost digit of the test score might attract more attention.[1] It suggests that individuals might perceive the ability signal provided by round scores to be significantly better than the ability signal provided by their preceding scores.

To examine the importance of round test scores, I implement a regression discontinuity design (RDD) around a round test score in the University Psychometric Entrance Test (PET), a high-stakes test required for university applicants in Israel. Specifically, I focus on the score cutoff of 600 in the first test for each individual. This setting is compelling since 600 is usually *just below* the admission threshold for enrolling in the most rewarding field in universities in Israel, computer sciences (CS).[2] Thus, individuals who score around 600 in their first PET face the dilemma of whether to invest effort in improving their admission-related outcomes (the PET score and high-school outcomes) to enroll in these rewarding university programs.

The identification assumption underlying my research is that the potential outcomes of test-takers are continuous at 600. This assumption is justified by the institutional context as manipulations of the cutoff are very unlikely, and admission chances for CS studies are not changing discontinuously at the cutoff. The empirical evidence also supports the identification assumption. Predetermined outcomes are continuous at the cutoff, and the probability of admission to CS degrees is very low and not discontinuously increasing at 600.

My empirical analysis reveals that the round test score significantly affects the decision to apply to CS. First, I study short-term applications, that is, within three years after taking the first test. I find that the rate of applications to CS degrees discontinuously increases by 1.4p.p. at the score cutoff of 600. Relative to a baseline rate of 3.8%, it reflects a 35% relative increase. Second, I study applications in the long term, up to twelve years after taking the test. I find that the increase in CS applications persists (1.7p.p.), reflecting a slightly smaller increase in relative terms (20%). Both estimates are statistically significant at the 99% level.

[1]Left-digit bias is humans' tendency to judge the difference between 600 and 599 to be larger than that between 601 and 600. Research in economics show that left-digit bias appears in consumers' decisions (Lacetera et al., 2012; Strulov-Shlain, 2021), and even in professional decision making (Olenski et al., 2020; Shurtz, 2022).

[2]The PET scores are between 200 and 800 (for more details, see Section 2).

Note that the first test scores of these individuals were below the admission threshold for CS degrees. Consistent with that, I show that individuals who score just above 600 improve their admission-related outcomes before applying to the program. They retake the PET and massively improve their scores.[3] They also improve their high school outcomes by achieving more credits, including more credits in CS programs in high school.

These results are consistent with the notion that individuals interpret the round score of 600 as a significant ability signal. To further examine this issue, I test for spillover effects on the younger siblings. Recent papers document that young siblings learn from the educational experiences of their older siblings (Gurantz et al., 2020; Altmejd et al., 2021). It might suggest that round test scores would affect younger siblings' decisions. Indeed, I show that younger siblings increase their rate of taking the PET within three years after their older sibling's test if the test score of the older sibling is just above 600.

I also document long-term gains for individuals who score just above 600 on their first PET. First, they increase their rates of enrollment and attainment of CS degrees. Second, they earn significantly more during the early years of their career. I argue that these findings are consistent with significant marginal returns to CS studies for individuals on the margin of *applying* to CS degrees in Israel.

Studying the heterogeneity of the effects reveals an intriguing pattern, where the increase in CS applications (and long-term outcomes, as well) is entirely driven by the younger test-takers.[4] I discuss two possible explanations for this finding. First, it might suggest that younger individuals are more sensitive to ability signals. An alternative explanation could be that test retaking is less costly for the younger test-takers, as they typically have a more extended period before applying to universities. In contrast, I do not detect significant heterogeneity of the effects by gender and socio-economic status ("SES").

Taken together, my results suggest that the round score of 600 was interpreted as a significant ability signal by the individuals in our sample, encouraged them to achieve more ambitious academic goals, and dramatically changed their career paths. These findings highlight the importance of understanding how young adults interpret ability signals, specifically by examining the role of heuristics and biases. The results in this paper could also have implications for a policy that aims to

[3]Note that the average effect of crossing 600 on PET retake is negative, consistent with recent evidence from the US, that individuals tend to retake the SAT less above round score cutoffs (Goodman et al., 2020). More important for my research is that, while the overall effect is negative, the impact on retaking and applying to CS is positive. I discuss explanations for these (seemingly contradictory) findings in subsection 5.2.

[4]There is a significant variation in ages among test-takers in Israel. Due to the mandatory army service, Jewish individuals often start their degrees at the ages of 21–23. Still, many individuals take their first PET during, or immediately after, high school (see Section 2).

increase participation in high-quality educational programs. For example, it might support a grading policy of rounding up individuals' test scores in some contexts, where ability signals may make a difference.

The evidence presented in this paper joins the growing literature on how ability signals could influence young adults' human capital decisions and labor market outcomes. From a theoretical point of view, when individuals are uncertain about their ability, ability signals could affect their choices (Manski, 1993; Altonji, 1993; Altonji et al., 2016). Empirical literature also supports this channel, with evidence that ability beliefs are a significant determinant of human capital decisions (e.g., Wiswall and Zafar, 2015).[5] More closely related, many papers show that human capital decisions are affected by test scores via the channel of learning about ability (Stinebrickner and Stinebrickner, 2012; Goodman, 2016; Arcidiacono et al., 2016; Avery et al., 2017, 2018; Bond et al., 2018; Graetz et al., 2020; Li and Xia, 2022; Tan, 2022). The contribution of my paper is to provide the first evidence (to the best of my knowledge) that young adults tend to follow the left-digit bias when they interpret ability signals and to show that this channel might have far-reaching consequences.

The results shown in this paper are also related to the literature on how test scores affect human capital and labor market outcomes via other channels that are (not necessarily) learning about own ability. Growing literature documents that these outcomes might be affected by exam luck that is caused by environmental factors (Ebenstein et al., 2016; Jisung, 2020), characteristics of the education system (Machin et al., 2020; Landaud et al., 2021), and teacher grading discretion (Diamond and Persson, 2016; Dee et al., 2019). My paper contributes to this literature by showing that the luck induced by informal cutoffs (specific round test scores) might have important implications.

More broadly, this paper is related to two additional branches of literature. First is the literature on the determinants of the field of study decisions (see Patnaik et al. (2020) for an updated review). Second is the literature on the behavioral aspects of educational choices, a relatively understudied topic (Lavecchia et al., 2016). A closely related paper is the paper by Goodman et al. (2020), which documents left-digit bias among students. The paper shows that American young adults tend to retake the SAT less above round score cutoffs. I document a similar pattern in Israel, showing that the PET retake rate decreases at the round score of 600. However, I contribute by showing that this is not the only effect of crossing the round score, at least in the Israeli context. Moreover, the impact I document here is in the opposite direction, as the ability signal induces individuals

[5]Recent papers have documented that students' ordinal rank significantly affects educational outcomes by enhancing students' ability beliefs (Pagani et al., 2021; Elsner et al., 2021).

to improve their academic aspirations. It also has long-term implications, with significant labor market gains.

The remainder of this paper is arranged as follows. In the following section (2) I provide background about the post-secondary education system in Israel; In section 3 I discuss the data and in section 4 I discuss the empirical strategy; In section 5 I show the estimation results of the impact of crossing 600 in the first PET on CS degrees and labor market outcomes; In section 7 I conclude.

2 Context and Background

The Israeli post-secondary education system is suitable for studying how young adults respond to round test scores. Applicants to academic institutions in Israel choose their fields of study before enrollment. They apply to universities after taking a standardized test, PET, with scores ranging between 200 and 800. Motivated by evidence on the left-digit bias, it is plausible that individuals would perceive round PET scores as a significantly better ability signal. I focus on the round score of 600 as it is just below the admission threshold for the most rewarding field in universities in Israel, CS. In this section, I provide details on the university fields in Israel and the admission requirements. I also provide details on both components of the admission requirements, the high school program and the PET.

2.1 University Fields of Study in Israel

Table C.1 shows the average outcomes of graduates of the most common fields of study at universities in Israel. The average annual earnings of CS graduates are the highest among these fields, consistent with prior evidence on the returns to the field of studies in Israel. For example, Achdut et al. (2019) document that CS is the most rewarding field of study in Israel in terms of earnings.[6] The second most rewarding field, as emerged from the table and the evidence presented in Achdut et al. (2019), is electrical engineering (EE).

Despite the high returns, many students do not attempt to participate in CS studies. For example, men apply for CS degrees in universities in Israel almost 200% more than women *with similar ability* (CBS data). Consistent with that, the table shows that the share of females is relatively low. It also shows that the average parental education is relatively high among graduates of CS degrees

[6] CS programs are also very challenging, and they attract highly talented individuals. For example, gifted children in Israel choose CS as their university field at very high rates (Lavy and Goldstein, 2022).

in universities. These findings motivate the study of how young adults in Israel decide whether or not to apply to CS degrees.

It is also important to note that Israel's academic system, during our sample period, includes seven universities and 50 colleges which offer undergraduate degrees.[7] Throughout the paper, I focus on university degrees for two reasons. First, university studies in Israel are of higher quality in terms of academic teaching, peers' ability, and other characteristics. Unsurprisingly, the labor market returns for a university degree are higher (Achdut et al., 2019). Thus, focusing on the decisions of young adults to pursue such degrees is of great interest in the Israeli context. Second, I observe only information on university applications, which is essential for this research as it allows me to study how individuals' *decisions* are affected.

2.2 Admission Decisions

The application process to academic institutions in Israel is as follows. Each individual is allowed to submit applications to as many institutions as she will. When applying to an academic institution, individuals choose their desired fields of study in advance (up to three fields of study, rated in preference order). Admission decisions are based mainly on two parameters, PET scores, and achievement in the Bagrut, the Israeli high school program. Admission decisions are determined by the summarized score ("Sechem"), a weighted average of the maximum PET score, and the mean composite Bagrut score. The weights and the exact method of calculation differ between institutions. The institution sets an official admission threshold for each program, the minimum Sechem required to enter the program. Some programs have additional requirements, such as a minimal PET threshold or Bagrut score in specific subjects.

Important for our research is the high official admission thresholds for CS degrees in Israeli universities. A PET score of 600 is usually below the threshold unless one has an extremely high Bagrut mean score. Unfortunately, complete data on the official admission thresholds for each program and year are not available (for details on the partial data I observe, see Appendix A.3). However, I observe the actual admission decisions made on each university application. These data allow me to empirically validate that admission chances for university CS degrees are low and smooth at 600.

Figure B.1 shows that the (unconditional) admission rates for CS degrees are very low at 600. With a *maximum* PET score of 600, the admission rate is less than 2%. For comparison, with a PET score of 700, the rate is above 10%.

[7]Two of the colleges became universities, after the sample period. Additionally, one university in Israel does not offer undergraduate degrees (only advanced degrees).

Second, Figure B.2 shows that the rate of admission for CS degrees, conditional on applying, is smooth at 600. The small (and insignificant) decrease could be due to the behavioral response documented throughout the paper. If the admission probability is not discontinuously changing at 600, and the rate of applications increases, we could expect that the additional applications would be from individuals with a relatively low probability of getting accepted. Thus, the admission probability could decrease. One way or another, the figure alleviates the concern that there is a discontinuous increase in the probability of admission to CS studies at 600.

Still, there is the caveat that although admission chances do not discontinuously increase at 600, individuals mistakenly believe they do. Indeed, young adults often have incomplete information about the admission criteria for academic programs. However, evidence suggests that individuals usually know that 600 is below the admission threshold for CS degrees. Figure B.3 shows that the vast majority of individuals with *first* PET scores of 600 who apply to CS degrees do this only after they retake the PET. About 80% of the individuals who score about 600 and still apply to CS retake the test before applying. In contrast, the rate of retaking among individuals who score about 700 and apply to CS is only 20%.[8]

2.3 Israel's High-School Program (Bagrut)

When entering high school (10th grade), students enroll in the academic or non-academic track. Students enrolled in the academic track receive a matriculation certificate ("Bagrut") if they pass a series of national exams in core and elective subjects taken between 10th and 12th grade. Students choose to take tests at various proficiency levels, each awarding one to five credit units per subject, depending on difficulty. Advanced level subjects award students more credit units (5 relative to 4 for an intermediate level and 3 for a basic level); A minimum of 20 credit units must qualify for a Bagrut certificate. About 52% of all high school seniors received a Bagrut in the 1999 and 2000 cohorts (Israel Ministry of Education, 2001).

2.4 The University Psychometric Entrance Test (PET)

The PET is administered by the National Institute of Testing and Evaluation ("NITE"). According to NITE, the PET is a tool for predicting academic success in higher education institutions in Israel. The test consists of three parts: quantitative reasoning, verbal reasoning, and English. The test scores in each domain range

[8]Note that I perform this analysis for EE studies and find very similar results, suggesting the probability of admission to EE degrees is low and smooth at 600, too.

from 50 to 150. The total score is a weighted average of the scores in all three domains (two-fifths weight for the quantitative and verbal reasoning scores each, and a fifth for the score in English). Finally, NITE normalizes the total scores to a score range of 200 and 800.

Figure B.4 shows the distribution of PET scores in our sample. The figure in panel a shows that the distribution is approximately normalized to a normal distribution. The 25^{th} and 75^{th} percentiles are 450 and 616. As expected, the density of observations is smooth at 600 in both samples. Note that the score cutoff of 600 is relatively high in the distribution of the scores, and only 22% of the individuals in our sample scored 600 or higher in their first tests.

3 Data

The analysis presented in this paper relies on an administrative database available from Israel's Central Bureau of Statistics (CBS). The data is based on merged datasets from multiple sources and includes information on all individuals enrolled in Israeli high schools (tenth grade) between 1995 and 2016. Two features make these data ideal for current research. First, I observe the universe of all PET undertaken by individuals in Israel. The data include the test score and the timing of each test. The second is a detailed description of applications to universities. The data includes the timing of application, the institution, the fields' ranking (1 to 3), and the admission decisions. These features allow me to investigate the relationship between test scores and field of study decisions. Other data features allow me to investigate long-term effects and heterogeneity by personal characteristics. Appendix A provide more details about the data.

3.1 Sample

The sample includes all individuals who participated in their *first* PET between 1995 and 2008. This sample definition allows me to study effects for at least twelve years after the test (data are available until 2020). The analysis is focused on the first test of each individual since not all individuals took a second test. Moreover, the ability signal is probably most significant in the first test of each individual. However, as a robustness test, I also study the effects of crossing 600 using an extended sample that includes all tests.

There is a significant variation in age at the first test among individuals in Israel. Figure B.5a shows that Arabs typically take the PET earlier than Jews. This difference is because Arabs enroll in post-secondary education younger, as they usually do not serve in the army.[9] Furthermore, Figure B.5b shows that Arabs tend to retake the PET much more than Jews. This gap is consistent with the fact that Arabs tend to have much lower PET scores than Jews on average (564 for Jews and 395 for Arabs). There is also variation among Jews, with about a half taking their first test at high school or two years after graduation (ages 20 and below) and about a half taking their first PET later.

3.2 Main Outcome Variables

University Applications. The primary outcome variables in this research capture field of study decisions. The primary interest is in the decision to apply to CS degrees in universities. For each individual in the sample, I define an indicator variable that takes the value one if and only if the individual applied for a CS university degree three years after the first PET. This outcome allows me to evaluate the short-term effect. I also define an indicator of applying to CS in any year, which helps me to distinguish between a short-term impact that dissipates with time and a persistent effect.

Additionally, I define indicator variables for university application to any field, any STEM field, and any non-STEM field to better understand the effects on alternative fields of study.[10] Lastly, I define an outcome variable that considers all fields and captures how ambitious the individual's applications are. "Predicted Income" represents the average income, twelve years after the test, among graduates of the program (field and institution) that the individual chose. In cases where the individual applied to more than one program, I use the maximal value.

University Degrees. First, I define an indicator variable for enrollment in a CS degree at a university. The enrollment data set includes missing information on the fields of study of each individual. However, when the university reports its admission decision, it also notes whether the individual started the degree or not. Thus, the applications' data set includes complete information on fields in which the individuals started their degrees. So, I use both data sets to construct the enrollment indicator, taking the value one if and only if the individual is marked as enrolled in a CS degree in the enrollment data set *or* as starting a CS degree in the applications' data set.

[9] Individuals start their compulsory military service in the Israeli army (IDF) immediately after high school graduation. Men serve for three years and women for two years. Therefore, most Jewish students start their post-secondary studies at ages 21-23.

[10] Note that I follow the Israeli CBS' broad definition for STEM fields, including Math, Statistics, Computer Sciences, Engineering, and Physical and Biological Sciences programs.

Second, I define an indicator variable for the attainment of a university CS degree. This indicator takes the value one if and only if the individual has completed a CS degree in a university (ever).

Degrees in Any Institution. While I do not observe information on college applications, I do observe information on the attainment of college degrees, including the field of study. Thus, I define also an indicator variable for completing a CS degree in any institution (university or college). I also explore the effects on the probability of attaining a BA degree (in any field and any academic institution) and on STEM and non-STEM degrees.

Labor Market Outcomes. I define labor market outcomes at the age of 28. I chose this age since this is the latest age in which I observe labor market information for all individuals in the sample. The labor market data is available until 2020, twelve years after my sample's last test. The youngest individuals are taking the PET at age 16; thus, twelve years after their test, they are at age 28. It implies that the labor market results represent early career stages immediately following university graduation.

First, I define an employment indicator with a value of 1 for individuals with a non-zero number of months of work and a non-zero income. Second, I define an indicator for self-employment. Third, I define an indicator variable for employment in the tech industry. According to the CBS definition, the tech industry includes multiple sectors classified as either tech manufacturing or tech services (Appendix A shows the complete list). Fourth, I construct income variables, focusing on total annual earnings. In the main specification, I use the natural log of the annual earnings, but I provide the results with many alternative definitions in the appendix. All income outcomes are measured in 2018 New Israeli Shekels (NIS).

I also use data on all individuals' labor market outcomes twelve years after the test. It is important to note that there is heterogeneity in the ages of individuals twelve years after the test (due to the heterogeneity in the ages at the first test). Thus, when examining the effects on earnings twelve years after the test, I use the annual earnings rank conditional on age.

3.3 Descriptive Evidence

Figure B.6 shows descriptive evidence of the relationship between first PET scores and applications to CS. The figure shows a clear positive association between the score and the probability of applying for a university degree in CS within three years after the test. This relationship could reflect learning from test scores. However, it could also reflect many other things, as the test score is probably correlated with the individuals' ability, SES, gender, admission chances, and other characteristics that might affect the decision to apply.

More importantly, the figures show a *discontinuous* increase in the rate of applications at 600. It suggests that young adults who score just above 600 interpret that as a positive ability signal, thus attempting to participate in CS degrees more. In Section 5, I study the discontinuous increase in applications at 600 more rigorously by implementing a regression discontinuity design. The figure also supports the decision to focus on the round score of 600. As expected, 600 is the most relevant round score cutoff when focusing on CS applications. This finding is not a surprise since this round score is usually just below the admission threshold.[11]

4 Empirical Strategy

What are the effects of crossing 600 in the first PET? In general, studying the effects of test scores is challenging since the score is correlated with many observed and unobserved individual characteristics. However, the left-digit bias implies a discontinuity in the perception of the test score at round scores. Thus, it allows the implementation of a regression discontinuity design (RDD) under the assumption that the potential outcomes are continuous at the cutoff. This assumption is justified in the context of the current research, as cutoff manipulations are implausible. Additionally, empirical evidence supports the validity of the identification assumption. First, I showed that the density of observations is smooth at 600 (subsection 2.4). Second, I showed that the admission chances for CS degrees do not increase above the cutoff (subsection 2.2). Third, I will show that all pre-determined outcomes are continuous at the cutoff (subsection 4.2).

4.1 Estimation.

To estimate the effects of crossing 600, define $Score_i$ as the first PET score of the individual i; $Above600_i$ as an indicator of crossing 600, that is $Score_i > 599$; R_i as a continuous variable that denotes the distance from the cutoff, $Score_i - 600$. With these definitions, I estimate the following equation using a sample of all

[11] One could ask why there is no significant discontinuous change in CS applications at other round scores. One natural explanation is that the lower cutoffs (400, 500) are too far from the admission thresholds. For individuals who achieve 400 or 500 in their first tests, it is hard to improve their test scores enough (such that they will have any chance to get accepted CS degrees). Regarding the higher cutoff, 700, the figure shows that the rate of CS applications is very high in the small window below the cutoff. It might suggest that individuals who get such test scores already have the confidence that they can study CS. Thus the ability signal could have less impact.

individuals who satisfy $Score_i \in [580, 619]$:

$$Y_i = \alpha_0 + \tau Above600_i + \alpha_1 R_i + \alpha_2 Above600_i R_i + \varepsilon_i^0 \tag{1}$$

Where τ is the coefficient of interest, which captures the impact of crossing 600 in the first PET on the outcome. All standard errors calculated throughout the analysis are heteroskedasticity-robust standard errors.

The choice of implementing a local linear RD approach within this bandwidth is somewhat arbitrary. I also estimate the effects on the primary outcomes using other specifications to validate that this choice does not drive the results. Specifically, I use the algorithm developed by Calonico et al. (2014) to estimate RD models with MSE optimal bandwidths. I run their algorithm with three different polynomial orders (0-2) to validate that the results are robust.

4.2 Falsification Tests

Under the identification assumption, pre-determined outcomes should be continuous at the cutoff. I test the continuity of pre-determined outcomes by estimating equation 1 with pre-determined outcomes as outcome variables. Panels a, b, and c of Table 1 show the results. In panel a, I use individual characteristics as outcome variables. These (pre-determined) outcomes are age at the test and dummies for females, Arabs, students in non-religious schools, and individuals who were born in Israel. In panel b, I use parental characteristics as outcome variables. This includes three dummies. The first is for individuals whose both parents were born in Israel. The second is for individuals whose sum of their parental income is above 250K NIS.[12] The third is a dummy for individuals whose parents have completed at least 13 years of schooling. Only one of the 10 estimates is statistically significant at the 90% level, and all estimated discontinuities are small in magnitude.

To further validate this issue, I predict the main outcome (CS applications) and the PET score using all covariates shown in the table and show graphical evidence that these predicted outcomes are continuous at 600 (Figure B.7). I also estimate the discontinuity in each of these predicted outcomes at the cutoff and show the estimates in Table 1.

[12] I use the average parental income when the individual in our sample is on ages 14, 15, and 16.

5 The Effects of Crossing 600 in the First PET

5.1 The Effects on University Applications

Figure 1a presents the main result of this paper, showing that the rate of CS applications increases discontinuously at the round score, 600, within three years after the test. Panel A of Table 2 shows the corresponding estimate (τ in equation 1) and its robust standard error. I find that the rate of applications to CS increases by 1.4p.p within three years after the test, implying a relative increase of 35%, which is statistically significant at the 99% level.

Furthermore, the table shows no significant change in the rate of applications for a university degree in any field within three years after the test. Interestingly, the increase in CS applications is not at the cost of a decrease in other STEM fields, as the impact on STEM applications is positive and statistically significant at the 95% level. Moreover, the rate of applications to EE, the following most rewarding field, also increases (the estimate is marginally significant at the 90% level). Consistent with that, I find a small decline in the rate of applications to non-STEM fields. I also find that the predicted annual income, based on the individual's applications, increases by about 6,180 NIS (approximately 1,500 USD). The increase is statistically significant at the 99% level.

Panel B of Table 2 shows the estimates for the impact on applications in the long term and their robust standard errors. The table shows that the increase in CS applications persists with a 1.7p.p increase, which implies a relative increase of about 20%. The estimate is statistically significant at the 99% level. The table also shows that the increase in EE applications dissipates with time, and the estimate becomes statistically insignificant. To summarize, the results show that individuals who score just above 600 apply to higher quality programs in terms of futural labor market outcomes, driven by a significant and persistent increase in the rate of applications to the most rewarding field, CS.

5.2 The Effects on Admission-related Outcomes

The main results show that individuals who score just above 600 in their first PET significantly increase their rate of applications to university CS degrees. However, 600 is usually below the admission threshold for university CS degrees. Hence, to meet the admission requirements, individuals should improve their Bagrut achievement and/or PET scores before applying. Thus, I study the impact of crossing 600 on these outcomes and report the results below.

Bagrut (High School) Achievement. Individuals might improve their Bagrut test scores to meet the admission requirements for CS. Note that about a third of the individuals in our sample are still in high school when taking their first PET; thus, they can change their Bagrut program after receiving their PET score. The rest, who took their PET after high-school graduation, can still improve their admission chances by retaking their Bagrut tests. Retaking the Bagrut tests after high school graduation is common among individuals wishing to enhance their post-secondary admission chances in Israel.

To examine the effects on Bagrut outcomes, I define two indicator variables. The first is an indicator for an extended high school program of more than 29 points. Recall that this is a relatively rich program, taken by only 25% of the individuals in our sample (the minimum for Bagrut eligibility is 20 points). The second indicator variable takes the value one if and only if the individual achieved 5 Bagrut credits in CS. Completing 5 CS credits in high school could reflect more directly the attempt to get into university CS degrees.

The results are shown in Figures 2a and 2b and in panel A of Table 3. I find a 1.8p.p. increase in the rate of 5 credits in Computer Sciences (10%, statistically significant at the 95% level) and a 2.0p.p. increase in the rate of achieving more than 29 credits (7%, significant at the 95% level). These results imply that individuals who score just above 600 in their first PET improve their Bagrut outcomes by achieving more scientific credits. When estimating the effect on the mean composite Bagrut score,[13] I do not detect a significant change. The point estimate is positive, 0.13, with a standard error of 0.201. It might be because the change in the Bagrut program is concentrated among a small number of individuals; thus, it is not strong enough to detect a change in the average Bagrut mean composite scores.

PET Retake. To explore the effects of crossing 600 on PET retake decisions, I define two indicator variables, taking the value one if and only if the individual retakes the test within three years and in any year. I also use the interaction of these variables with the indicator for CS applications to capture the impact of crossing 600 on retaking the PET *and* applying to CS degrees. I estimate equation 1 with each of these outcomes.

The results are shown in Figures 2c and 2d and in panel B of Table 3. First, I find that the average effect of crossing 600 on the retake rate is negative. This finding is consistent with evidence from the US, where individuals tend to retake the SAT less above round scores (Goodman et al., 2020). What could explain this

[13]The method of calculation is slightly different when applying to different academic institutions. I followed the method used for the Hebrew University, as described here.

pattern? Young adults' expectations or goals may be consistent with the left-digit bias; thus, getting a score of 600 is good enough for some individuals, and they decide to avoid retaking if they achieve this goal. This channel could negatively affect the retake rate since individuals who cross 600 would be less motivated to improve their test scores.

However, when estimating the effects on retaking *and* applying to CS, I find positive effects consistent with the idea that most individuals applying to this program (with first PET scores of around 600) retake the test before applying. The increase in CS applications and retaking is 1.1p.p., which indicates that about 65% of the increase in CS applications is by individuals who retook the test before applying. This result further validates that the increase in applications at 600 is not due to a change in (the real or perceived) admission chances.

To better understand retaking patterns, I define indicator variables for achieving a (maximum) PET score above 630 and 640. These outcomes help distinguish between individuals who retake the PET and improve it little (less than 630) and individuals who massively improve their PET score such that they become closer to the admission thresholds for CS degrees. Interestingly, I do not detect a statistically significant change in the rate of individuals who achieve 630 or more. It suggests that the negative effect on retake rate is pronounced among individuals with lower academic aspirations (probably not those who would apply to CS degrees in any case). In contrast, when interacting this indicator with the indicator for CS applications, I find an increase of about 1p.p. (significant at the 95% level). The results are similar when using even the higher threshold of 640, suggesting that a non-negligible number of individuals improve their test scores before applying to CS degrees.

In summary, I find evidence for both mechanisms of improving PET scores and improving Bagrut outcomes before applying to CS degrees in universities. These findings are consistent with the institutional background information that a PET score of 600 is usually below the admission thresholds for CS. They also suggest that individuals invest effort and time in their attempt to participate in CS degrees as a response to the (heuristic) ability signal.

5.3 Spillovers to Younger Siblings

Recent literature has documented that individuals' educational decisions are affected by the experiences of their older siblings (Joensen and Nielsen, 2018; Qureshi, 2018; Gurantz et al., 2020; Altmejd et al., 2021). Thus, to further support the channel of learning about ability from round test scores, I also examine the effects of crossing 600 on the testing decisions of younger siblings. To examine that, I define an outcome variable that takes the value one if and only if the individual has any young sibling who took a PET within three years after the test (of the older sibling, this is the individual in our sample). I also use an indicator for a positive number of young siblings as a balance check.

Figure 3 shows the results, where there is no discontinuous change in the rate of individuals with young siblings at 600, but there is a (marginally) significant increase in the testing rate among younger siblings (2.0p.p., 8.5% relative change). Panels a and b of Table 4 show the estimates and also the impact on younger siblings testing decisions ever (positive but statistically insignificant). The next panel, c, shows the effects on alternative outcome variables (the share of younger siblings tested). These results support the learning channel, and they suggest that young siblings also learn about their potential from the round test scores of their older siblings.

5.4 Robustness of the Main Results

MSE-optimal bandwidths estimates. The results so far were based on a local linear RD estimation within a fixed bandwidth (580 to 619). I also report the main results using MSE optimal bandwidths with three different polynomial orders (0-2) and with two choices of the kernel function (uniform and triangular) to validate that the results are robust. The results are shown in Table 5. The table shows that the main results reported above are highly robust to both choices of the polynomial order, the bandwidth, and the kernel function.

Using all tests in the estimation. Additionally, I validate that the results presented above are not sensitive to the choice to focus on the first tests only. As a robustness test, I extend the sample to include all tests taken between 1995 and 2008 and find similar results (see Table C.2).

Placebo exercise at non-rounded cutoffs. I also implement a placebo exercise, estimating the discontinuous change in the main outcomes at alternative, non-rounded, score cutoffs. I find desired null results, as shown in Table C.3.

5.5 Heterogeneity of the Effects

How does the increase in CS applications at 600 vary by personal characteristics? First, I would like to separate the Jews and the Arabs, as we saw that the testing regularities of these two groups are different. Additionally, I split the sample of Jewish test-takers into two by their age during the test (20 and below or older). This split is motivated by the differences between Jewish individuals who take the test early and those taking the test later. For example, the rate of individuals with 5 Bagrut credits in CS is 25% among the younger test-takers and only 10% among the older. Similarly, the average mean composite Bagrut score is 100.6 for the younger and 92 for the older. This evidence shows that the younger test-takers are positively selected.

Panel A of Table 6 shows the estimated effects (τ in equation 1) on applications to CS, among each group, separately.[14] I find that the increase in applications is the largest among the Arab test-takers. However, the sample of Arabs is quite small (less than 2,000 effective observations); thus, it has serious power limitations. Among Jews, the effect is entirely driven by the younger test-takers, with no change in the rate of applications among the older.

These results motivate examining the heterogeneity by age among Jewish test-takers. Thus, for each individual i, I define Age_i as the age at the first test and Age^*_i as the years difference relative to the age 18 ($Age^*_i = Age_i - 18$).[15] Then, I estimate the following equation:

$$\begin{aligned} Y_i = &\beta_0 + \beta_1 Age^*_i + \delta Above600_i + \delta^a Age^*_i \cdot Above600_i + \\ &\beta_2 R_i + \beta_3 Age^*_i \cdot R_i + \beta_4 Above600_i \cdot R_i + \\ &\beta_5 Age^*_i \cdot Above600_i \cdot R_i + \varepsilon^1_i \end{aligned} \tag{2}$$

Where the coefficient δ represents the effect on those at age 18 when taking their first tests, and δ^a represents how the effect varies by age.

Figure 4 show the estimated impact of crossing 600 on CS applications. The figures show that the impact is statistically significant for the younger test-takers (20 and below) and insignificant for the older. The estimate for the slope is statistically significant in both cases, suggesting that the effects among Jews are decreasing by age. Why are the effects more pronounced among the younger test-takers? First, it could be that younger individuals are more sensitive to ability signals. Second, recall that the PET score of 600 is below the admission threshold for CS degrees; thus, retaking is necessary if the individual decides to apply to such a program. Intuitively, the costs of retaking are much lower for younger individuals, as they are typically not employed and have more years

[14]Table C.4 shows the falsification tests on each group separately. The table shows that the pre-determined outcomes are continuous among each group at 600, with only 4 of 36 estimates that are significant at the 90% level (similar to what we expect to get by chance).

[15]Note that I have binned this variable to take care of outliers, such that the lowest age becomes 17 and the highest becomes 23. However, I validate that this binning procedure does not affect the results.

before enrolling in post-secondary education.[16] Consistent with that, the baseline test retake rates are 22% for older test-takers and 52% for younger ones.

Motivated by this finding, in what follows, I focus on the younger Jewish test-takers (age 20 or below). I prefer to focus on this sample since there are very few Arabs in the sample, and they significantly differ in their (observed and unobserved) characteristics from the Jewish individuals in the sample. I exclude the older Jewish individuals as the analysis revealed that they are not affected by crossing the round score cutoff. Still, I also report the estimation results for the samples of older Jews, and Arabs, in the Appendix.

Next, I study the heterogeneity of the effects by gender. I define $female_i$ as an indicator for females and estimate a version of equation 2 with $female_i$ replacing Age_i^*. Panel A of Table C.5 shows the results. The estimated effect of crossing 600 on CS applications is stronger among males in the long term. However, the difference between the effects is statistically insignificant. Moreover, the baseline mean in the outcomes is much larger for males (the difference is significant at the 99% level), which reflects the much higher tendency of males to apply to CS degrees. Thus, the relative increase is similar for males and females (about 20%).

I also examine how the effects vary by SES, by estimating a version of equation 2 with $High_i$ replacing Age_i^*. $High_i$ is an indicator variable that takes the value one if and only if both parents of individual i completed more than 12 years of schooling. Panel B of Table C.5 shows that the increase in CS applications at the round score cutoff of 600 is slightly more significant for high-SES individuals. One explanation for why the impact would have been more pronounced among the high SES individuals is that they may be less sensitive to paying the cost associated with PET retaking. However, the difference between the estimates is insignificant, possibly due to power limitations.

It is interesting to relate these findings to the growing literature on heterogeneity in response to ability signals. There is evidence that females are more sensitive to ability signals relative to men, in some contexts (Ahn et al., 2019; Owen, 2021a; Coffman et al., 2021; McEwan et al., 2021; Franco, 2019). In other contexts, there are similar effects for females and males (Bestenbostel, 2021; Owen, 2021b). Additionally, there is evidence that lower SES individuals are more affected by ability signals in some contexts (Graetz et al., 2020; Avery and Goodman, 2022). In the empirical context of my paper, both males and females and individuals with low and high SES were similarly affected by the ability signal. However, I show that the effects vary by age, with young individuals being the main group affected.

[16] Cost, in this context, could reflect both the monetary cost and the time and effort required for test retake.

6 The Effects of Crossing 600 on Long Term Outcomes

The finding that individuals increase their rate of CS applications at the score cutoff of 600 suggests that individuals' human capital investment decisions are affected by heuristic interpretation of ability signals. An open question is whether these decisions affect long-term human capital *outcomes*. It is also an open question what are the labor market implications. In this subsection, I attempt to answer these questions by estimating the effects of crossing 600 in the first PET on CS degrees, employment, and earnings during the early career stages.

6.1 Effects on Degrees.

Figure 5 shows the effects on enrollment in, and on the attainment of, CS degrees using the sample of younger test-takers.[17] Panel A of Table 7 presents the estimates of the impact of crossing 600 on these outcomes. I find a 1.8p.p. increase (33%) in the rate of enrollment in CS degrees, which is significant at the 99% level. Estimating the impact on CS degrees in universities, I find an increase of about 1.0p.p. (29%). However, this result is only marginally statistically significant due to power limitation at the 90% level. These results suggest that the increase in CS applications is by individuals who get into CS programs and complete these programs in a probability that is not different relative to the baseline probability. Importantly, for some individuals, the ability signal of the round test score significantly affects their human capital *outcomes*, as they study more towards CS degrees in universities.

When estimating the effects on the attainment of CS degrees in any institution (including colleges), I find a statistically significant increase that is larger in absolute terms and reflects a similar relative increase of 25%. This result suggests that the increase in enrollment in CS degrees in universities is not due to a decrease in enrollment in CS degrees in colleges. Lastly, I study the effects on degrees in other fields and find that the increase in CS degrees is not due to a decrease in other STEM fields but in the rate of non-STEM degrees (Table C.8). The table also shows that the average effect of attaining a BA degree in any institution and field is zero.

[17] I also report the results using the sample of older Jewish test-takers (Table C.6) and Arabs (Table C.7). Unsurprisingly, there are no significant long-term effects for the older test-takers. For the Arabs, the estimates are noisy due to the small sample size.

6.2 Effects on Labor Market Outcomes.

Figure 6 and Panel B of Table 7 shows the estimated effects on labor market outcomes at age 28. There is no significant change in the employment rate, and the increases in employment in the tech industry and self-employment are statistically insignificant.[18] However, I find a 9% increase in the annual earnings, which is significant at the 95% level.[19] The age of 28 is relatively young. Thus, the results are restricted to the very early stages of the career. However, I also analyze the effects on labor market outcomes 12 years after the test. Still, these results refer to relatively early career stages, but the ages range between 28 and 32. The results are very similar, with a significant increase in the earnings rank, conditional on age (Panel C).[20] Interestingly, there is a significant (95%) increase in the rate of self-employment twelve years after the test (19%).

The estimated effects on annual earnings might imply huge local labor market returns. According to our estimation, there is about a 1.8p.p. increase in the enrollment rate in CS degrees. Taken at face value, the estimated impact of 9% on annual income implies huge labor market returns of about 500% (that is, 9/0.018) for these individuals. A more conservative approach is to consider the increase in the rate of CS applications, as it is possible that individuals who applied but did not get CS degrees still got into other high-quality university programs. Taking this conservative approach still implies huge labor market returns of about 360% (that is, 9/0.025). This finding is essential, as it suggests that the returns to CS degrees in universities in Israel, among individuals who are at the margin of *applying* to CS degrees, are enormous. Thus, a policy that would affect the field of study choices might have considerable benefits.

Still, this measure should be taken with caution, as crossing the round score might directly affect annual earnings, not through the impact on education.[21] One such direct effect could be since individuals are psychologically affected by the ability signal, thus taking other ambitious career (or human capital) choices. Second, it is possible that employers are following the left-digit bias such that the increase in earnings reflects the labor market signaling value of the test scores, too. However, As documented earlier, most individuals in our sample retake the PET. Thus, their first PET scores are less relevant, and it seems unlikely that

[18]The tech industry is the leading industry for graduates of CS degrees (and STEM degrees in general) in Israel. For a further detailed discussion, see Lavy and Goldstein (2022).

[19]The figures also show a negative trend within the bandwidth. Note that our research design is neutral regarding these slope estimates. Notably, the discontinuous increase at 600 is robust to the specification, as I show in subsection 6.3.

[20]Conditioning on age is essential when analyzing the effects on earnings in a fixed period after the test since individuals in our sample are taking their test at different ages.

[21]An intuitive way to think about it is within the framework of instrumental variable (IV), where the outcome is earnings, the endogenous variable is CS studies, and the IV is the indicator *Above*600. Within this framework, it is plausible that the exclusion restriction is violated.

this channel drives the increase in annual earnings. Third, some individuals in our sample may still be working part-time jobs, due to the early stage of their careers (e.g., if they pursue a Master's degree). Still, it is unlikely that all these other channels would drive a massive increase of 300-500% in annual earnings. Thus, a reasonable interpretation is that the increase in earnings reflects high labor market returns for marginal individuals.

6.3 Robustness of the Long Term Results

MSE-optimal bandwidths estimates. The long-term results presented above are based on a local linear RD estimation within a fixed bandwidth (580 to 619). Here, I show that the main results are very robust to specification by using MSE optimal bandwidths with three different polynomial orders (0-2) and two choices for the kernel function (uniform and triangular). The results are shown in Table 8. The table shows that the long-term results reported above are highly robust to both choices of the polynomial order, the bandwidth, and the kernel function.

Placebo exercise at non-rounded cutoffs. I also implement a placebo exercise, estimating the discontinuous change in the main outcomes at alternative, non-rounded, score cutoffs. I find desired null results, as shown in Table C.9.

Alternative outcomes. In addition, Table C.8 shows the effects on alternative measures of the annual earnings, which are the average annual earnings (including and excluding zeros), the average annual earnings from salaried employment (including zeros), and the (natural log of the) average annual earnings at age 29. The table also shows that crossing the round score does not significantly affect marriage decisions.

6.4 Heterogeneity of the Long Term Results

To explore how the long-term effects vary by age, I estimate equation 2 using the full sample of Jewish individuals. As expected, Figure 7 shows that all effects decrease with age, consistent with the findings in Section 5.

Studying heterogeneity of the long-term impact by gender, I find that the effects (on the rate of attaining CS degrees and on the earnings rank) are smaller in magnitude for females (Panel A of Table C.10). However, due to power limitations, these differences in the estimated effects for males and females are statistically insignificant. Panel B of the table shows that the effects are pronounced among high and low-SES individuals, with no significant difference between the effects on both groups.

7 Conclusion

In this paper, I examined the importance of round test scores. I showed that crossing a round score, 600, in the first PET positively affected the rate of applications to CS, one of the most rewarding fields of study in Israel. Moreover, I showed that individuals who score just above 600 retake the test and improve their test scores before applying to CS as well as improve their high school outcomes, which are both ways to improve admission chances. Importantly, I also showed that crossing the round score of 600 induced significant long-term implications. Individuals who score just above 600 in their first PET study more for CS degrees in universities, and earn more. All the effects documented in the paper are driven by the younger test-takers. In general, the effects are similar for men and women, and for low and high-SES test-takers.

Taken together, these results indicate a significant career change for individuals who score just above 600 in their first PET. It suggests that an uninformative signal might have far-reaching consequences on human capital investment and on labor market outcomes. An additional piece of results that supports the interpretation of the round test score being a meaningful signal is that crossing 600 affects the testing decisions of the younger siblings, too.

The results presented in this paper are important for two main reasons. First, studying how ability signals affects human capital investment is central to understanding how young adults make these decisions. My research contributes a novel finding to this literature by showing that learning might occur heuristically. Second, this research could potentially provide lessons for a policy that aims to increase participation in high-quality education programs. For example, it may motivate a very cheap affirmative action with potentially large benefits: to round up the test scores. It is difficult to predict in advance the magnitude of the impact of such policies in different contexts. However, the costs of implementing them are very low, making them a promising investment.

References

Achdut, Lea, Elad Guttman, Noam Zussman, Idan Lipiner, and Inbal Ma'ayan, "Returns to education in universities and colleges," Technical Report 2019.

Ahn, Thomas, Peter Arcidiacono, Amy Hopson, and James Thomas, "Equilibrium Grade Inflation with Implications for Female Interest in STEM Majors," Technical Report 2019.

Altmejd, Adam, Andres Barrios-Fernandez, 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," *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), 48–83.

_ , Peter Arcidiacono, and Arnaud Maurel, "The Analysis of Field Choice in College and Graduate School: Determinants and Wage Effects," *Handbook of the Economics of Education*, 2016, *5*, 305–396.

Arcidiacono, Peter, Esteban Aucejo, Arnaud Maurel, and Tyler Ransom, "College Attrition and the Dynamics of Information Revelation," Technical Report 2016.

Avery, Christopher and Joshua Goodman, "Ability Signals and Rigorous Coursework: Evidence from AP Calculus Participation," Technical Report 2022.

_ , Oded Gurantz, Michael Hurwitz, and Jonathan Smith, "Giving College Credit Where It Is Due: Advanced Placement Exam Scores and College Outcomes," *Journal of Labor Economics*, 2017, *35* (1), 67–147.

_ , _ , _ , and _ , "Shifting College Majors in Response to Advanced Placement Exam Scores," *The 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," Technical Report 2021.

Bond, Timothy N., George Bulman, Xiaoxiao Li, and Jonathan Snith, "Updating Human Capital Decisions: Evidence from SAT Score Shocks and College Applications," *Journal of Labor Economics*, 2018, *36* (3), 807—-839.

Calonico, Sebastian, Matias D. Cattaneo, and Rocio Titiunik, "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs," *Econometrica*, 2014, *82* (6), 2295–2326.

Coffman, Katherine B., Paola Ugalde Araya, and Basit Zafar, "A (Dynamic) Investigation of Stereotypes, Belief-Updating, and Behavior," Technical Report 2021.

Dee, Thomas S, Will Dobbie, Brian A Jacob, and Jonah Rockoff, "The causes and consequences of test score manipulation: Evidence from the New York regents examinations," *American Economic Journal: Applied Economics*, 2019, *11* (3), 382–423.

Diamond, Rebecca and Petra Persson, "The long-term consequences of teacher discretion in grading of high-stakes tests," Technical Report, National Bureau of Economic Research 2016.

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

Franco, Catalina, "How does relative performance feedback affect beliefs and academic decisions? Evidence from a field experiment," Technical Report 2019.

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.

Goodman, Serena, "Learning from the Test: Raising Selective College Enrollment by Providing Information," *The Review of Economics and Statistics*, 2016, *98* (4), 671—-684.

Graetz, Georg, Björn Öckert, and Oskar Nordström Skans, "Family Background and the Responses to Higher SAT Scores," Technical Report 2020.

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.

Jisung, Park R, "Hot Temperature and High Stakes Performance," *Journal of Human Resources*, 2020.

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

Lacetera, Nicola, Devin G. Pope, and Justin R. Sydnor, "Heuristic Thinking and Limited Attention in the Car Market," *American Economic Review*, 2012, *102* (5), 2206–36.

Landaud, Fanny, Éric Maurin, Barton Willage, and Alexander Willén, "Getting Lucky: The Long-Term Consequences of Exam Luck," Technical Report 2021.

Lavecchia, A.M., H. Liu, and Philippe Oreopoulos, "Behavioral Economics of Education: Progress and Possibilities," *Handbook of the Economics of Education*, 2016, *5*, 1–74.

Lavy, Victor and Yoav Goldstein, "Gifted Children Programs' Short and Long-Term Impact: Higher Education, Earnings, and the Knowledge-Economy," Technical Report 2022.

Li, Hongyan and Xing Xia, "Grades as Signals of Comparative Advantage: How Letter Grades Affect Major Choices," *Available at SSRN*, 2022.

Machin, Stephen, Sandra McNally, and Jenifer Ruiz-Valenzuela, "Entry through the narrow door: The costs of just failing high stakes exams," *Journal of Public Economics*, 2020, *190*, 104224.

Manski, Charles F., "Schooling as Experimentation: A Reappraisal of the Postsecondary Dropout Phenomenon," *Economics of Education review*, 1993, *8* (4), 305–312.

McEwan, Patrick J., Sherin Rogers, and Akila Weerapana, "Grade Sensitivity and the Economics Major at a Women's College," *AEA Papers and Proceedings*, 2021, *111*, 102–106.

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

Owen, Stephanie, "College Field Specialization and Beliefs about Relative Performance," Technical Report 2021.

_ , "Do Grade Signals Drive the Gender Gap in STEM? Evidence From a Regression Discontinuity," Technical Report 2021.

Pagani, Laura, Simona Comi, and Federica Origo, "The effect of school rank on personality traits," *Journal of Human Resources*, 2021, *56* (4), 1187–1225.

Patnaik, Arpita, Mathhew Wiswall, and Basit Zafar, "College Majors," Technical Report 2020.

Qureshi, Javaeria A, "Additional returns to investing in girls' education: Impact on younger sibling human capital," *The Economic Journal*, 2018, *128* (616), 3285–3319.

Shurtz, Ity, "Heuristic thinking in the workplace: Evidence from primary care," *Health Economics*, 2022.

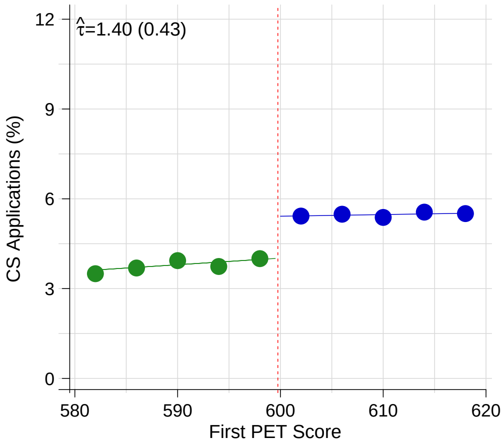
Stinebrickner, Todd and Ralph Stinebrickner, "Learning about Academic Ability and the College Dropout Decision," *Journal of Labor Economics*, 2012, *30* (4), 707—-748.

Strulov-Shlain, Avner, "More than a Penny's Worth: Left-Digit Bias and Firm Pricing," *Chicago Booth Research Paper*, 2021, (19-22).

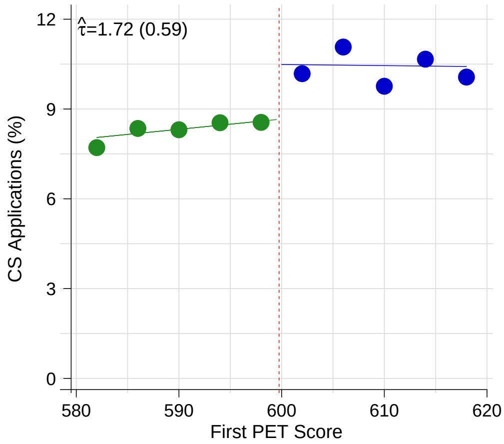
Tan, Brandon Joel, "The Consequences of Letter Grades on Labor Market Outcomes and Student Behavior," *Journal of Labor Economics*, 2022, pp. –.

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.

Figure 1: The Impact of Crossing 600 on Applications to Computer Sciences



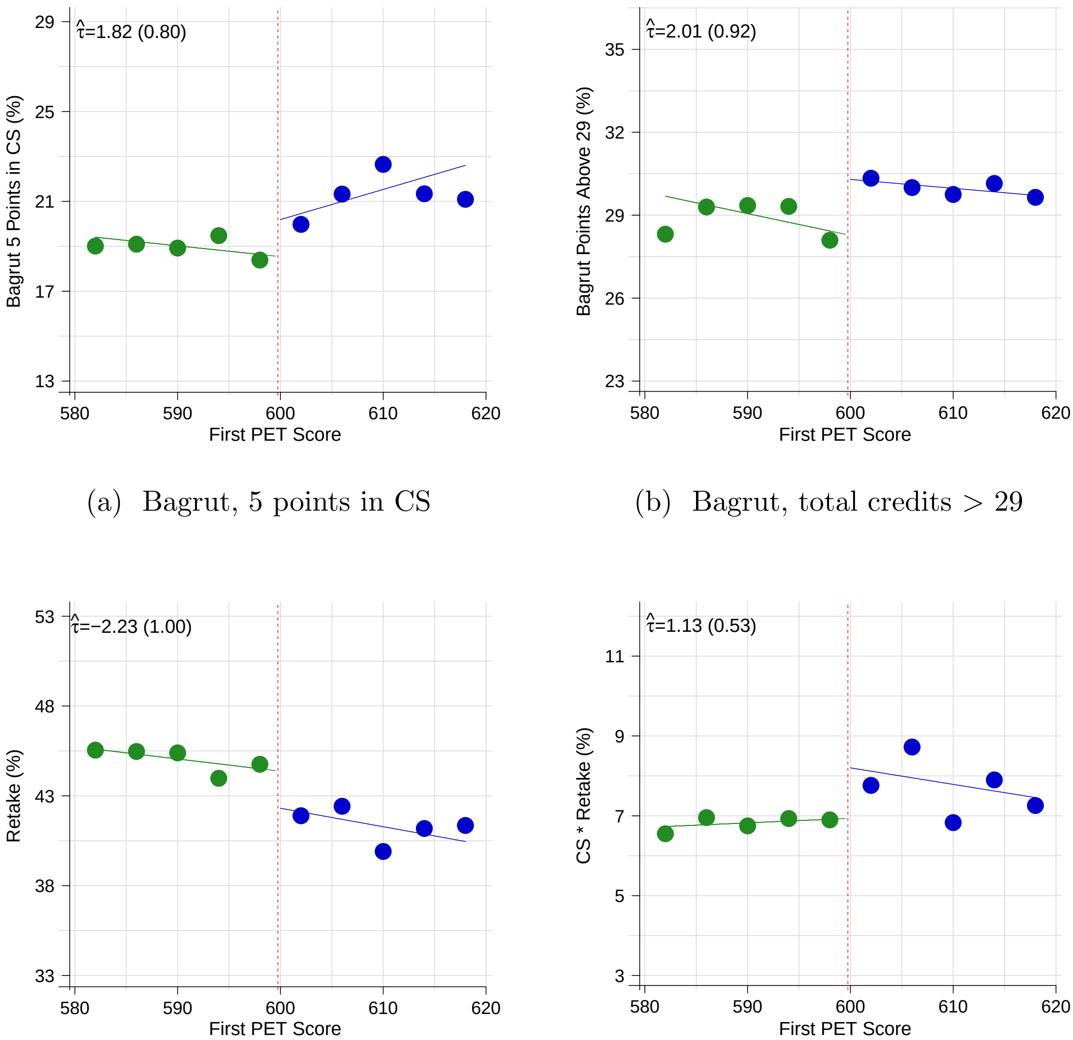
(a) Within three years



(b) Ever

Notes: The figures show the impact of crossing 600 in the first PET on applications to CS degrees in universities. The y-axes show rate of applications within three years after the first test (a) and ever (b). The x-axis shows the total score in the first PET (in bins of four points). The sample includes all individuals in Israel who participated in their first PET during 1995-2008 (43,911 effective observations). The red dashed vertical line represents the score cutoff of 600. The blue and green solid lines are based on the estimation of equation 1. The figures also report the estimates (and robust standard errors) for the coefficient of interest, τ, which represents the impact of crossing 600 on the outcome variable.

Figure 2: The Impact of Crossing 600 on Admission-related Outcomes



(a) Bagrut, 5 points in CS

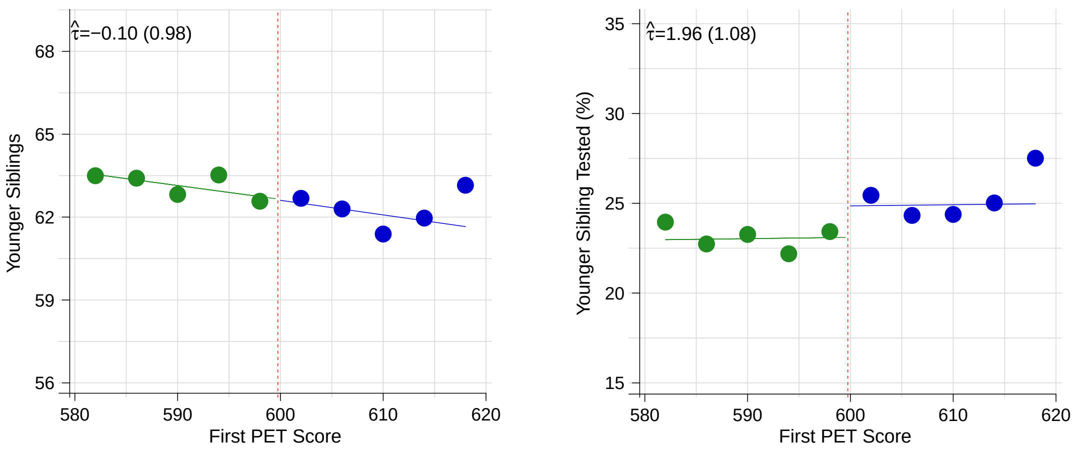
(b) Bagrut, total credits > 29

(c) PET retake

(d) PET retake $\times$ CS application

Notes: The figures show the impact of crossing 600 in the first PET on admission-related outcomes. The y-axes show the rate of achieving five credits in CS in the Bagrut (a), the rate of achieving a total number of 29 credits or more in the Bagrut (b), the rate of PET retake (c), and the rate of PET retake *and* applying to CS degrees in universities (d). The x-axis shows the total score in the first PET (in bins of four points). The sample includes all individuals in Israel who participated in their first PET during 1995-2008 (43,911 effective observations). The red dashed vertical line represents the score cutoff of 600. The blue and green solid lines are based on the estimation of equation 1. The figures also report the estimates (and robust standard errors) for the coefficient of interest, τ, which represents the impact of crossing 600 on the outcome variable.

Figure 3: The Impact of Crossing 600 on Younger Siblings' Testing Decisions

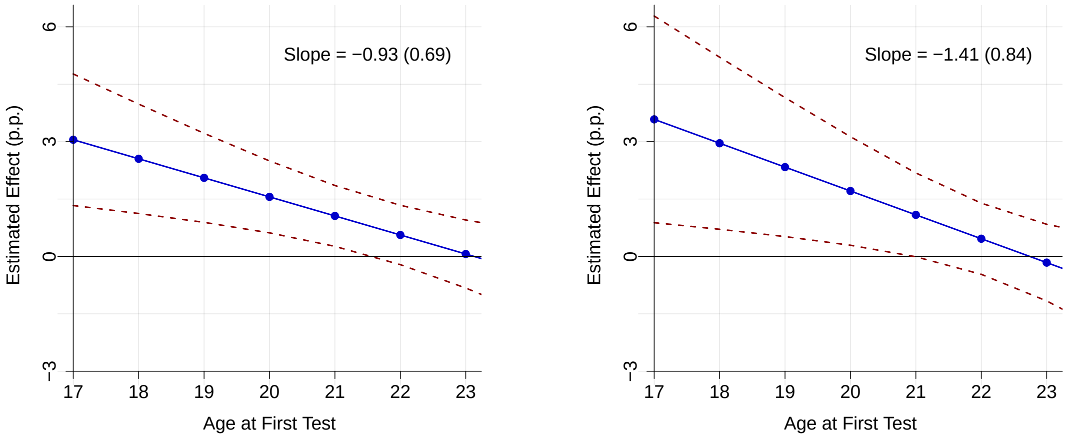


(a) Has siblings (%) – **Balance check**

(b) Has siblings taking the PET (%)

Notes: The figures show the impact of crossing 600 in the first PET on younger siblings' testing decisions. The y-axes show the rate of individuals in our sample with younger siblings (a) and the rate of testing among younger siblings of the individuals in our sample within three years after the test. The x-axis shows the total score in the first PET (in bins of four points). The sample includes all individuals in Israel who participated in their first PET during 1995-2008 (43,911 effective observations). The sample in panel b is restricted to individuals with at least one sibling (27,860 effective observations). The red dashed vertical line represents the score cutoff of 600. The blue and green solid lines are based on the estimation of equation 1. The figures also report the estimates (and robust standard errors) for the coefficient of interest, τ, which represents the impact of crossing 600 on the outcome variable.

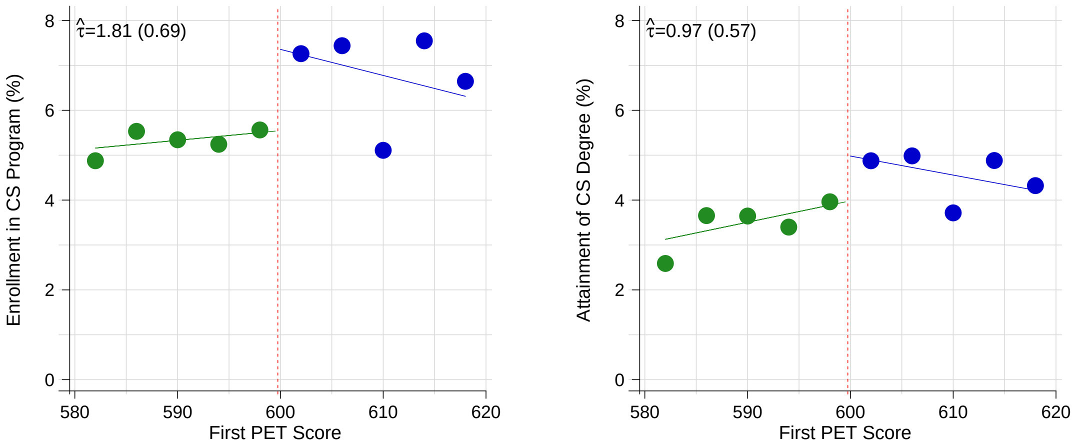
Figure 4: Heterogeneous Effects of Crossing 600 on CS Applications, by Age



(a) Within three years (b) Ever

Notes: The figures show the estimated effects (solid blue line) and their 95% confidence intervals (red dashed lines) of crossing 600 in the first PET, by the age at the first test. The estimates are based on the estimation of equation 2, calculated by $\delta + \delta^a \cdot (Age - 18)$ with CS applications within three years (a) and ever (b) as the outcome variables. The numbers represent percentage points. The sample includes all Jewish individuals in Israel who participated in their first PET during 1995-2008 and got a test score within the window of 580 to 620 (42,019 observations). The age variable is binned such that 17 (23) represents 17 and younger (23 and older).

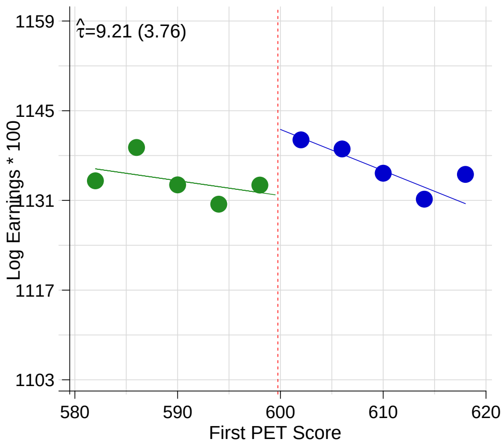
Figure 5: The Impact of Crossing 600 on Computer Sciences Degrees, Sample of Younger Test-Takers (Ages 20 and below)



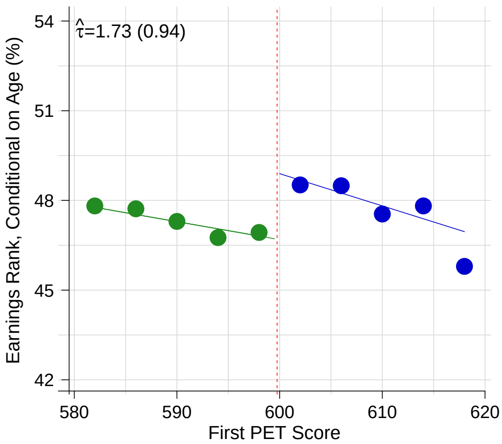
(a) Enrollment (b) Attainment

Notes: The figures show the impact of crossing 600 in the first PET on CS degrees. The y-axes show the rate of enrollment in CS degrees in universities (a), and the rate of attainment of CS degrees in universities (b). The x-axis shows the total score in the first PET (in bins of four points). The sample includes all Jewish individuals in Israel who participated in their first PET during 1995-2008 when they are at age 20 or below (22,963 effective observations). The red dashed vertical line represents the score cutoff of 600. The blue and green solid lines are based on the estimation of equation 1. The figures also report the estimates (and robust standard errors) for the coefficient of interest, τ, which represents the impact of crossing 600 on the outcome variable.

Figure 6: The Impact of Crossing 600 on Annual Earnings, Sample of Younger Test-Takers (Ages 20 and below)



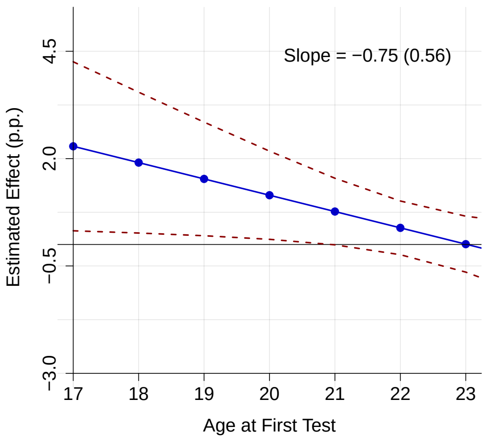
(a) Log earnings × 100



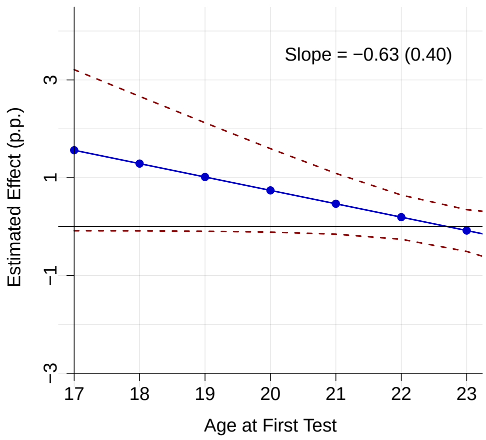
(b) Earnings rank, conditional on age

Notes: The figures show the impact of crossing 600 in the first PET on annual earnings. The y-axes show the natural log of the annual earnings at the age of 28 (a) multiply by 100, such that the numbers represent percentages, and the earnings rank, twelve years after the test, conditional on age (b). The x-axis shows the total score in the first PET (in bins of four points). The sample includes all Jewish individuals in Israel who participated in their first PET during 1995-2008 when they are at age 20 or below (22,963 effective observations). The red dashed vertical line represents the score cutoff of 600. The blue and green solid lines are based on the estimation of equation 1. The figures also report the estimates (and robust standard errors) for the coefficient of interest, τ, which represents the impact of crossing 600 on the outcome variable.

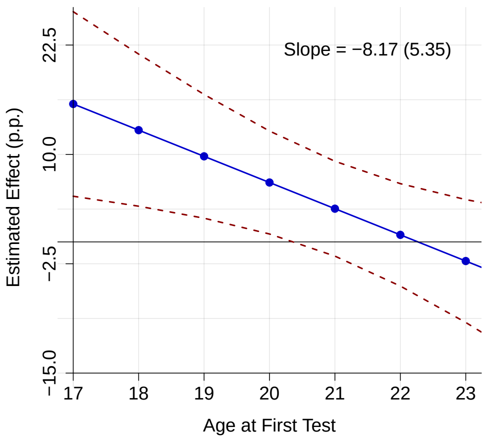
Figure 7: Heterogeneous Effects of Crossing 600 on Long Term Outcomes, by Age



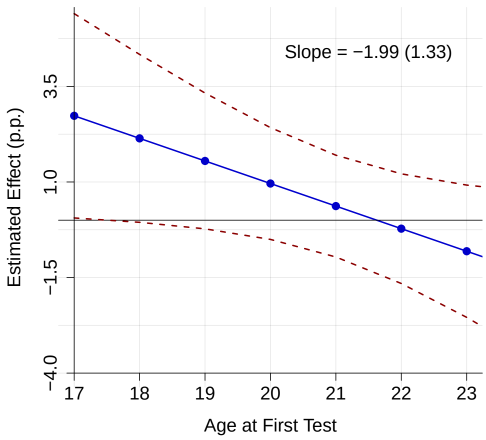
(a) CS Enrollment



(b) CS Degree



(c) Log earnings × 100



(d) Earnings rank, conditional on age

Notes: The figures show the estimated effects (blue solid line) and their 95% confidence intervals (red dashed lines) of crossing 600 in the first PET, by the age at the first test. The estimates are based on the estimation of equation 2, calculated by $\delta + \delta^a \cdot (Age - 18)$, using the following outcomes: enrollment in CS degrees in universities (a), attainment of CS degrees in universities (b), log annual earnings (c), and earnings rank (d). The numbers represent percentage points in all panels except c, where numbers represent parentage's. The sample includes all Jewish individuals in Israel who participated in their first PET during 1995-2008 and got a test score within the window of 580 to 620 (42,019 observations). The age variable is binned such that 17 (23) represents 17 and younger (23 and older).

Table 1: The "Impact" of Crossing 600 on Pre-determined Outcomes

	Baseline Mean	Estimate
	(1)	(2)
A. Individual		
Age	19.82	−0.050
		(0.050)
Female (%)	55.70	−1.805*
		(1.001)
Arab (%)	4.780	0.572
		(0.411)
Non-religious school (%)	82.63	0.253
		(0.771)
Born in Israel (%)	83.96	−0.786
		(0.740)
B. Parents		
Born in Israel (%)	42.88	1.124
		(0.999)
Income > 250K NIS (%)	42.33	−1.072
		(0.999)
Post high school education (%)	50.32	0.474
		(1.008)
C. Test		
Year	2003.41	−0.068
		(0.069)
Month	7.37	−0.066
		(0.070)
D. Predicted values		
Predicted CS (%)	8.06	0.128
		(0.130)
Predicted PET score	553.25	−0.411
		(0.868)
Number of observations		43,911

$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Notes: The table shows the estimated discontinuous change in pre-determined outcomes at the score cutoff of 600 in the first PET. Columns (1) and (2) show the baseline mean and the estimates for τ in equation 1 (with robust standard errors in parenthesis). The baseline sample includes all individuals in Israel who participated in their first PET during 1995-2008. The effective sample is restricted to individuals with first test scores within the window of 580 and 619 (43,911 observations).

Table 2: The Impact of Crossing 600 on University Applications

Outcome	Baseline Mean (1)	Estimate (2)
A. Within three years		
Computer Sciences (%)	3.78	1.397***
		(0.427)
Electrical Engineering (%)	3.21	0.854**
		(0.374)
Any (%)	40.30	0.708
		(0.998)
STEM (%)	19.63	2.176***
		(0.835)
Non STEM (%)	20.67	-1.468^{*}
		(0.823)
Predicted income (10,000 NIS)	22.58	0.608***
		(0.229)
B. Ever		
Computer Sciences (%)	8.27	1.719***
		(0.59)
Electrical Engineering (%)	6.47	0.426
		(0.503)
Any (%)	67.92	-1.333
		(0.920)
STEM (%)	36.50	1.641*
		(0.982)
Non-STEM (%)	31.41	-2.974^{***}
		(0.937)
Predicted income (10,000 NIS)	23.66	0.391**
		(0.179)
Number of observations		$43,911$

$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Notes: The table shows the estimated impact of crossing the score cutoff of 600 in the first PET on university applications. Columns (1) and (2) show the baseline mean and the estimates for τ in equation 1 (robust standard errors in parenthesis). The baseline sample includes all individuals in Israel who participated in their first PET during 1995-2008. The effective sample is restricted to individuals with first test scores within the window of 580 and 619 (43,911 effective observations).

Table 3: The Impact of Crossing 600 on Admission-related Outcomes

	Baseline Mean	Estimate
	(1)	(2)
A. Bagrut		
5 credits CS (%)	18.94	1.817** (0.800)
Total credits > 29 (%)	28.83	2.014** (0.920)
Mean composite score	97.26	0.134 (0.201)
B. PET retake		
Within three years (%)	28.78	−1.397 (0.907)
Ever (%)	45.06	−2.227** (0.999)
Three years × CS (%)	4.29	1.104*** (0.425)
Ever × CS (%)	6.80	1.131** (0.526)
C. Maximum PET score		
> 630 (%)	31.77	−1.018 (0.947)
> 640 (%)	27.93	−0.564 (0.920)
> 630 × CS (%)	5.56	0.998** (0.486)
> 640 × CS (%)	5.13	0.808* (0.467)
Number of observations		43,911

*$p < 0.1$; **$p < 0.05$; ***$p < 0.01$

Notes: The table shows the estimated impact of crossing the score cutoff of 600 in the first PET on admission related outcomes (PET retake and maximal scores, and high-school outcomes). Columns (1) and (2) show the baseline mean and the estimates for τ in equation 1 (robust standard errors in parenthesis). The baseline sample includes all individuals in Israel who participated in their first PET during 1995-2008. The effective sample is restricted to individuals with first test scores within the window of 580 and 619 (43,911 observations).

Table 4: The Impact of Crossing 600 on Younger Siblings' Testing Decisions

	Baseline Mean	Estimate
	(1)	(2)
A. Balance test		
Number of siblings	1.112	−0.008
		(0.024)
Has siblings (%)	63.13	−0.098
		(0.975)
B. Has siblings taking the PET		
Within three years (%)	23.19	1.961*
		(1.085)
Ever (%)	67.26	0.829
		(1.188)
C. Share of siblings taking the PET		
Within three years (%)	15.01	1.412*
		(0.805)
Ever (%)	53.81	0.818
		(1.097)
Number of observations		43, 911

$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Notes: The table shows the estimated impact of crossing the score cutoff of 600 in the first PET on younger siblings' testing decisions. Columns (1) and (2) show the baseline mean and the estimates for τ in equation 1 (robust standard errors in parenthesis). The baseline sample includes all individuals in Israel who participated in their first PET during 1995-2008. The effective sample in Panel a is restricted to individuals with first test scores within the window of 580 and 619 (43,911 observations). The effective sample in Panels b, and c is also restricted to individuals with at least one sibling (27,860 individuals).

Table 5: The Impact of Crossing 600 on Outcomes using Optimal MSE Bandwidths

	0-order	Linear	Quadratic
	(1)	(2)	(3)
A. Triangular kernel			
CS applications, three Years (%)	1.369***	1.482***	1.477***
	(0.386)	(0.451)	(0.513)
	[13]	[30]	[47]
CS applications, ever (%)	1.748***	1.801***	1.723**
	(0.545)	(0.594)	(0.705)
	[11]	[36]	[49]
CS × Retake (%)	1.059**	1.158**	1.109*
	(0.426)	(0.516)	(0.616)
	[14]	[34]	[48]
Bagrut 5 credits in CS (%)	1.690**	2.407***	2.017*
	(0.707)	(0.916)	(1.100)
	[12]	[17]	[29]
Has siblings taking the PET (%)	2.126***	1.995*	1.942
	(0.779)	(1.054)	(1.204)
	[18]	[30]	[47]
B. Uniform kernel			
CS applications, three years (%)	1.280***	1.507***	1.462***
	(0.350)	(0.443)	(0.524)
	[8]	[31]	[34]
CS applications, ever (%)	1.842***	1.858***	1.795**
	(0.521)	(0.590)	(0.710)
	[7]	[33]	[34]
CS × Retake (%)	0.986**	1.221**	1.227*
	(0.421)	(0.508)	(0.628)
	[10]	[30]	[34]
Bagrut 5 credits in CS (%)	1.490**	2.245***	2.287**
	(0.716)	(0.850)	(1.056)
	[7]	[17]	[24]
Has siblings taking the PET (%)	2.032**	2.167*	1.716
	(0.880)	(1.106)	(1.266)
	[10]	[19]	[34]

$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Notes: The table shows the estimated effects of crossing 600 in the first PET on outcomes, using MSE optimal bandwidths based on the algorithm developed by Calonico et al. (2014). The baseline sample includes all individuals in Israel who participated in their first PET during 1995-2008. The upper panel (A) shows the results using a uniform kernel, and the lower panel (B) shows the results using a triangular kernel. Columns (1), (2), and (3) show bias-corrected estimates (and robust standard errors) for different polynomial fit orders. The optimal bandwidths chosen are shown in squared parenthesis.

Table 6: The Impact of Crossing 600 on CS Applications, by Groups

	Arabs		Jews, Age≤20		Jews, Age>20	
	Mean	Estimate	Mean	Estimate	Mean	Estimate
	(1)	(2)	(3)	(4)	(5)	(6)
Within three years (%)	17.34	5.784	3.72	2.172***	2.33	−0.137
		(3.638)		(0.632)		(0.491)
Ever (%)	24.12	10.982***	11.09	2.590***	3.00	−0.432
		(4.197)		(0.946)		(0.556)
Number of observations	1,892		22,963		19,054	

$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Notes: The table shows the estimated impact of crossing the score cutoff of 600 in the first PET on outcomes among different groups of test-takers. Columns (1), (3) and (5) show the baseline mean of the outcome. Columns (2), (4), and (6) shows the estimates for τ in equation 1 (and their robust standard errors). The baseline sample includes all individuals in Israel who participated in their first PET during 1995-2008. The effective sample is restricted to individuals with first test scores within the window of 580 and 619. The sample in columns (1) and (2) includes only Arab test-takers (1,892 observations). The sample in columns (3) and (4) includes only Jewish test-takers, who were at age 20 or below when taking their first test (22,963 observations). The sample in columns (5) and (6) includes only Jewish test-takers, who were at age 21 or above when taking their first test (19,054 observations).

Table 7: The Impact of Crossing 600 on Long-term Outcomes, Sample of Younger Test-Takers (Ages 20 and below)

	Baseline Mean	Estimate
	(1)	(2)
A. CS degrees		
Enrollment (%)	5.30	1.809***
		(0.686)
Attainment (%)	3.44	0.972*
		(0.57)
Attainment, including college degrees (%)	5.77	1.430**
		(0.701)
B. Labor market, age 28		
Log Annual Earnings $\times$ 100	1133.98	9.208**
		(3.758)
Employment (%)	87.02	-0.265
		(1.183)
Self employment (%)	5.73	0.827
		(0.650)
Employment in the tech industry (%)	13.86	0.932
		(1.015)
C. Labor market, 12 years after the test		
Earnings rank, conditional on age (%)	47.31	1.731*
		(0.945)
Employment (%)	90.58	0.875
		(1.364)
Self employment (%)	8.89	1.638**
		(0.821)
Employment in the tech industry (%)	12.61	0.465
		(1.008)
Number of observations		22,963

$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Notes: The table shows the estimated impact of crossing the score cutoff of 600 in the first PET on long-term outcomes. Columns (1) and (2) show the baseline mean and the estimates for τ in equation 1 (robust standard errors in parenthesis). The sample is based on all individuals in Israel who participated in their first PET during 1995-2008 when they were at age 20 or below, and it includes only individuals with first test scores within the window of 580 and 619.

Table 8: The Impact of Crossing 600 on Long-term Outcomes using Optimal MSE Bandwidths, Sample of Younger Test-Takers (Ages 20 and below)

	0-order	Linear	Quadratic
	(1)	(2)	(3)
A. Triangular kernel			
CS enrollment (%)	1.626***	1.731**	1.965**
	(0.612)	(0.701)	(0.869)
	[11]	[39]	[32]
CS degree (%)	0.815	0.904	0.905
	(0.523)	(0.599)	(0.679)
	[10]	[30]	[38]
Log earnings × 100	8.974**	10.311***	9.712**
	(3.531)	(3.914)	(4.648)
	[13]	[32]	[39]
Earnings rank	1.583*	1.836**	1.793
	(0.837)	(0.93)	(1.149)
	[15]	[31]	[41]
B. Uniform Kernel			
CS enrollment (%)	1.507***	1.875**	1.902**
	(0.575)	(0.729)	(0.854)
	[8]	[24]	[29]
CS degree (%)	0.773*	0.814	1.011
	(0.451)	(0.625)	(0.672)
	[8]	[20]	[39]
Log earnings × 100	7.193**	11.829***	10.754**
	(2.957)	(4.026)	(4.186)
	[11]	[26]	[42]
Earnings rank	1.610**	2.151**	1.874*
	(0.786)	(0.982)	(1.086)
	[10]	[22]	[40]

$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Notes: The table shows the estimated effects of crossing 600 in the first PET on outcomes, using MSE optimal bandwidths based on the algorithm developed by Calonico et al. (2014). The baseline sample includes all individuals in Israel who participated in their first PET during 1995-2008 when they were at ages 20 an below. The upper panel (A) shows the results using a uniform kernel, and the lower panel (B) shows the results using a triangular kernel. Columns (1), (2), and (3) show bias-corrected estimates (and robust standard errors) for different polynomial fit orders. The optimal bandwidths chosen are shown in squared parenthesis.

Online Appendix

Appendix A Data Appendix

A.1 Data Sources

I use an administrative database from the Central Bureau of Statistics of Israel (CBS), which allows restricted access to this data in their protected research lab. The data include merged datasets from multiple sources. CBS matched and merged the datasets using the individual-level national ID number, and the matching is perfect without missing observations. Data are available for all enrollees in Israeli high schools (tenth grade) between 1995 and 2016.

The merged data file used for this research combines information from the following datasets: 1) National Institution for Testing and Evaluation provides information on the University Psychometric Entrance Test (PET); It includes the scores and timing of all tests ever taken by each individual in the sample. 2) Higher Council of Education records of post-secondary completed degrees, the institution of study (colleges and universities), the field of study (one or two), and completion year; essential additional features of these data are the applications and admission decisions information. I observe these data for universities in all years and colleges in 2008 and later only. Thus, here I use only university applications. 3) Israel Tax Authority (ITA) provides data on the earnings of employees and self-employed individuals from 2000-2018 and a three-digit code of industry of employment. 4) The population registry data includes a fictitious individual national ID number that appears in all the data sets described below and enables the matching and merging of the files at the personal level. It also contains information on the following student's family background variables: birth year, sex, locality, number of siblings, country of birth, and parental countries of birth. 5) Ministry of Education has provided administrative data on Israeli high schools' universe since 1995. It provides data on student's study program by subject and level, a variety of high school achievement measures, and test scores in all national matriculation exams in 10th-12th grades.

A.2 Sample and Definitions

The main sample includes all individuals in Israel who participated in their first PET between 1995 and 2008. In the main text, I explain why the such restriction is crucial. The definitions for the main outcome variables used in this research are described in the main text. Here, I provide supplementary information.

Field of Study. The fields of study appear in both data sets (degrees and applications) in a three-digit code, following the CBS classification (CBS website).

I followed the CBS in identifying Computer Sciences (Electrical Engineering) degrees and applications according to code 900 (1020). Additionally, I used the CBS definitions to classify degrees and applications to STEM and non-STEM fields.

Institution Type. I followed the CBS definition for institution types. All degrees and applications were classified as belonging to either an elite university, a non-elite university, or a college. I could not identify specific institutions.

Admission Decisions and Enrollment. In the application data set, I observe the admission decisions for each application. The institution also reports whether the individual started the program. Unfortunately, in cases where the individual applied to more than two fields in her application, I do not know for sure for each of the fields she got accepted. I observe an indicator for getting admitted to the first rank and one of the two other ranks. However, most applications to Computer Sciences and Electrical Engineering are either stand-alone or combined with other high-quality STEM programs.

Tech Industry. The definition of the tech industry is based on working in services or manufacturing tech companies. According to the CBS definition, these include the following industries: Pharmaceutical products for human and veterinary uses, Office and accounting machinery and computers, Electronic components, Electronic communication equipment, Industrial equipment for control and supervision, medical and scientific equipment, Aircraft (manufacturing); Telecommunications, Computer, and related services, Research and Development (services).

Bagrut Test Scores. For individuals in our sample who graduated high-school before 2004, I observe grouped Bagrut test scores. I used the median group values to construct the test score variable. For example, if the individual's score is between 91 and 100, I assign her the value of 95.5.

Parental Education and Income There is a small number of missing values in the data on parental education and income. In cases of missing values, I used the partner's education income to impute a value. If both parent had missing values, I assigned the average value in my sample.

Predicted Outcomes. I have predicted the main outcomes (CS and EE applications within three years) by fitting a logistic regression of the pre-determined variables on the outcome. I used the full sample of first tests (any test score, any age, including Arabs) for the estimation. The following explanatory variables are included: Gender, Age at test, Indicator for Arabs, Indicator for individuals who were born in Israel, Indicator for Individuals that both their parents were born in Israel, Indicator for regular school, parental income, parental education, number of siblings, and the year and the month of the test.

A.3 Information on Official Admission Thresholds

Admission decisions for most post-secondary programs in Israel are based on the summarized score ("Sechem"), which is a weighted average of the maximum PET score and the mean composite Bagrut score. The weights and the exact method of calculation differ between institutions. The institution sets an official admission threshold for each program, which is the minimum Sechem required to enter the program. Official information on university admission thresholds for each field of study is unavailable for the sample period. However, I could find partial data on the thresholds in specific institutions in specific years. For example, I have data on admission thresholds at Tel Aviv University in 2012, Hebrew University during 1995-2007, Haifa University in 2012, and Ben Gurion University in 2013. In all cases, the minimal Sechem (weighted average of the Bagrut mean composite score and the PET score) required for getting into CS programs in universities is very high, making it almost impossible to get into these programs using the PET score of 600. This is consistent with the empirical evidence I provide in Section 5, where individuals retake the PET (and improve their Bagrut outcomes) before applying. It is also consistent with the observation that the conditional acceptance rate in my data is continuous above 600.

For example, for a candidate with a PET score of 600, the minimal mean Bagrut score required to enter CS program at Tel Aviv University in 2012 was 114, which is very high and rare (the top percentile among individuals with first PET scores of about 600).[22] Indeed, the requirements in other universities are very similar in any case where I have the official data. It is important to note one particular issue. The official admission requirements for all scientific programs (including CS) in Ben Gurion University in 2013 include a Sechem threshold and an additional PET threshold of 600. However, this requirement does not bind in most cases of CS applications since only a few individuals have Bagrut mean composite scores of about 110, which is the required score to meet the admission

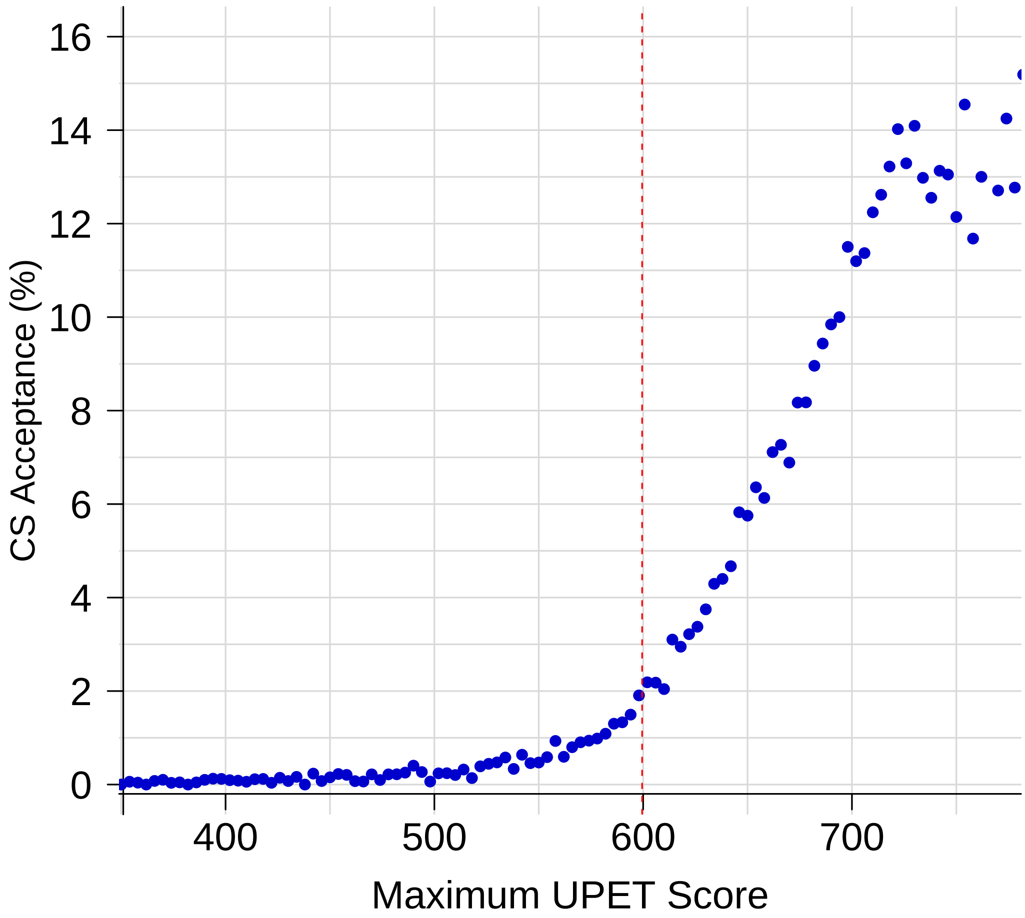
[22]These numbers are larger than 100 due to the calculation method of the mean composite score. These methods give extra 10-30 points for test scores in some high school programs (e.g., programs with five credits).

requirements with a PET score of 600. Another condition to getting into the CS programs, in some cases, is participating in the 5 or 4 credits Bagrut program in math (and achieving a minimum score).

It is also important to note that there are many nondirect ways to get into university programs in Israel. For example, pre-academic programs help students improve their admission chances. In addition, students who complete one year of university studies may sometimes switch fields. However, I am unaware of any case where PET score of 600 plays a role in the possibility of getting into a CS program in a university.

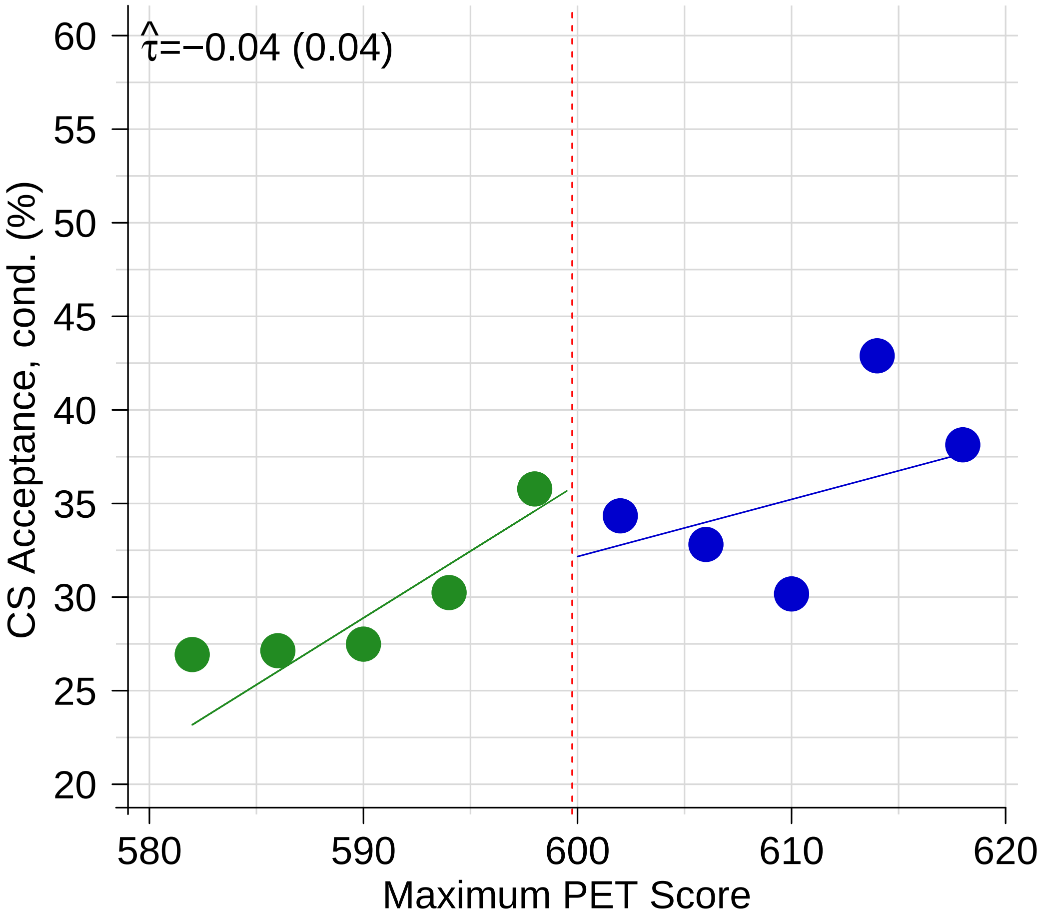
Appendix B Figures

Figure B.1: Unconditional Rate of Admission to Computer Sciences Programs in Universities



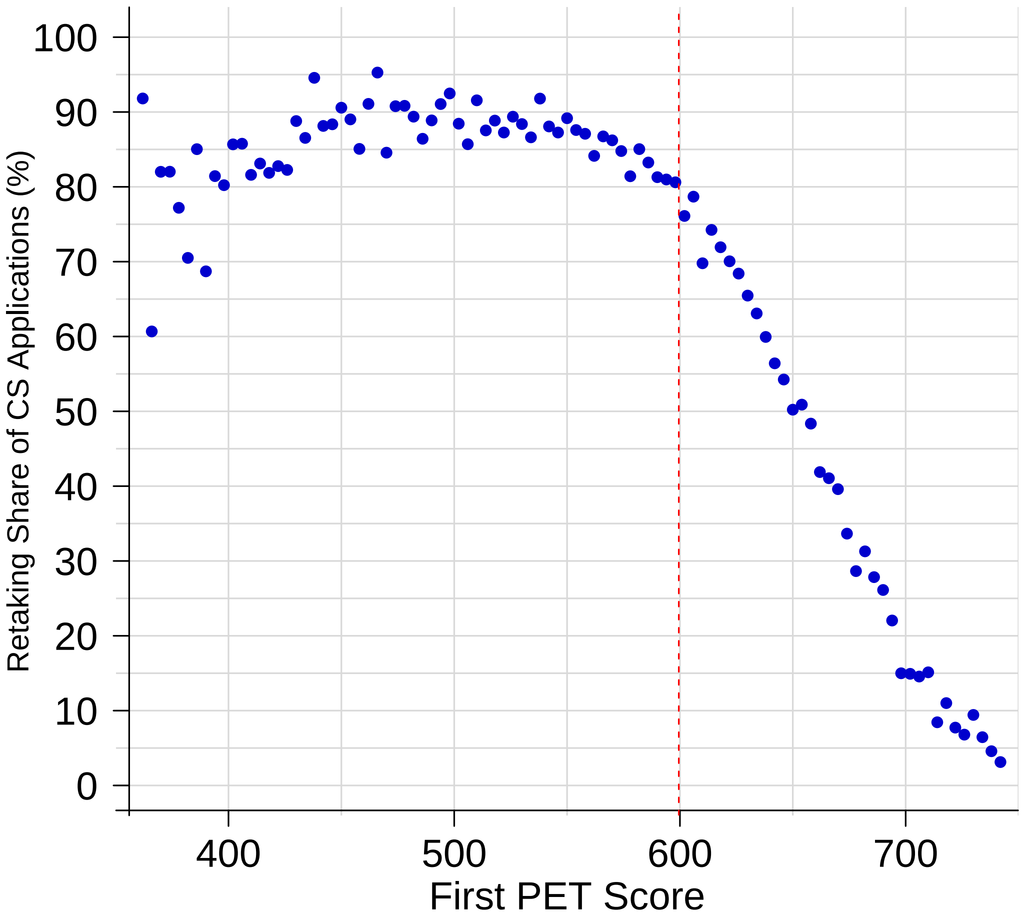
Notes: The figure shows the relationship between the rate of acceptance to CS programs in universities in Israel (y-axis) and the maximum PET scores (in bins of four points) at the individual level (x-axis). The sample includes all individuals in Israel who participated in their first PET during 1995-2008 (332,947 observations observations). The red dashed vertical line represents the score cutoff of 600.

Figure B.2: Rate of Admission to Computer Sciences Programs in Universities, Conditional on Applying



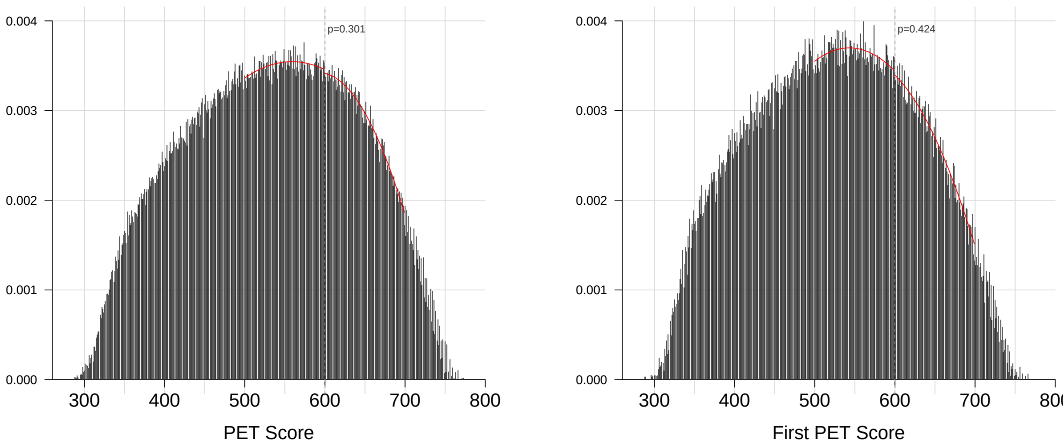
Notes: The figure shows the estimated discontinuity in the rate of acceptance to CS programs at the score cutoff of 600 in the first PET. The y-axis show rate of acceptance to CS programs (conditional on applying). The x-axis shows the first PET score (in bins of four points). The sample includes all individuals in Israel who participated in their first PET during 1995-2008 (43,911 effective observations). The red dashed vertical line represents the score cutoff of 600. The blue and green solid lines are based on the estimation of equation 1. The figure also reports the estimates (and robust standard errors) for the coefficient of interest, τ, which represents the impact of crossing 600 on the outcome variable.

Figure B.3: PET Retaking Share Among CS Applicants, by First PET Score



Notes: The figures show the relationship between the share of PET retake and first PET scores, among individuals who apply to CS programs. The y-axis show the share of individuals who retook the PET. The x-axis shows the first PET score (in bins of four points). The sample includes all individuals in Israel from cohorts 1979 and later, who participated in their first PET during 1995-2008 and apply to Computer Sciences (332,947 observations). The red dashed vertical line represents the score cutoff of 600.

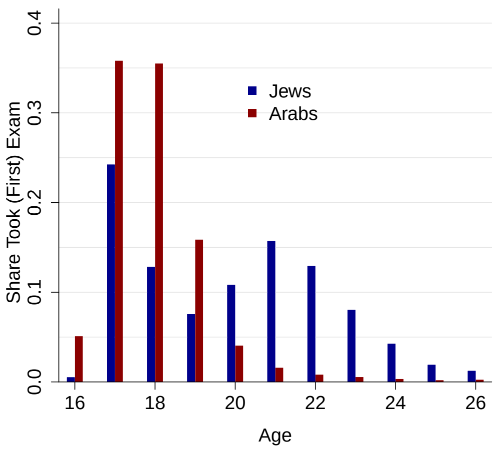
Figure B.4: PET Scores Distribution



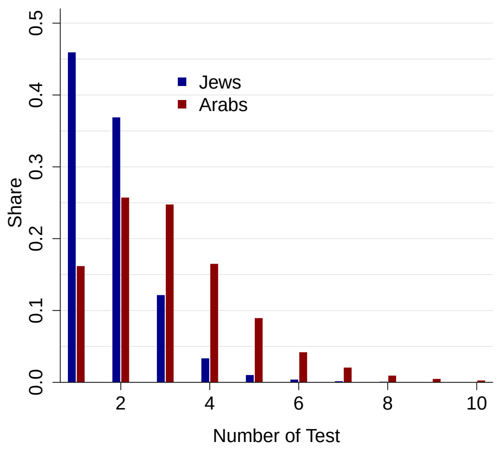
(a) All Tests (b) First Tests

Notes: The figures plot the share of observations (y-axis) according to the PET score (x-axis). The left panel include all tests, and the right panel includes only the first test of each individual. The figures show also the estimated densities on each side of the cutoff, 600, by allowing quadratic fit on each side of the cutoff (within a bandwidth of 100 points on each side), and the p-value for the discontinuity at the cutoff (0.285 and 0.418). The sample includes all PET taken in Israel during 1995-2008 (550,162 observations in panel a and 332,947 observations in panel b).

Figure B.5: Testing in the PET, by Population Group



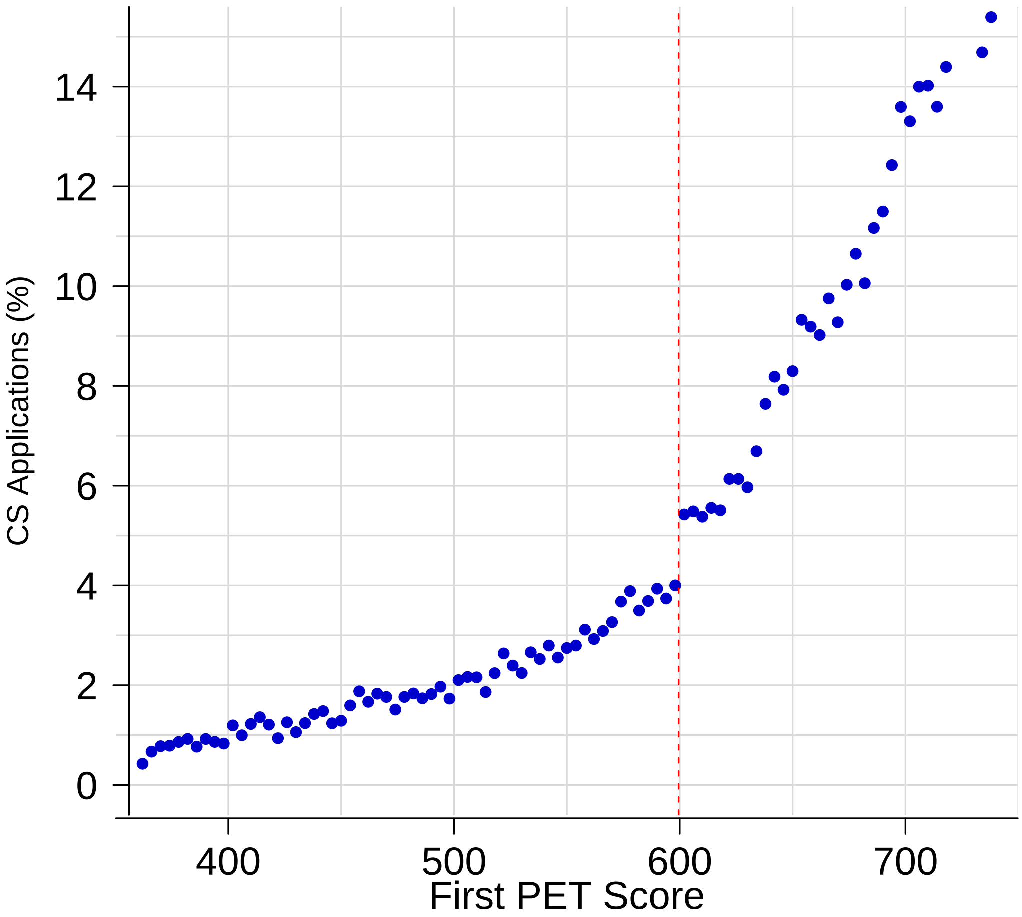
(a) Age at First Test



(b) Number of Tests

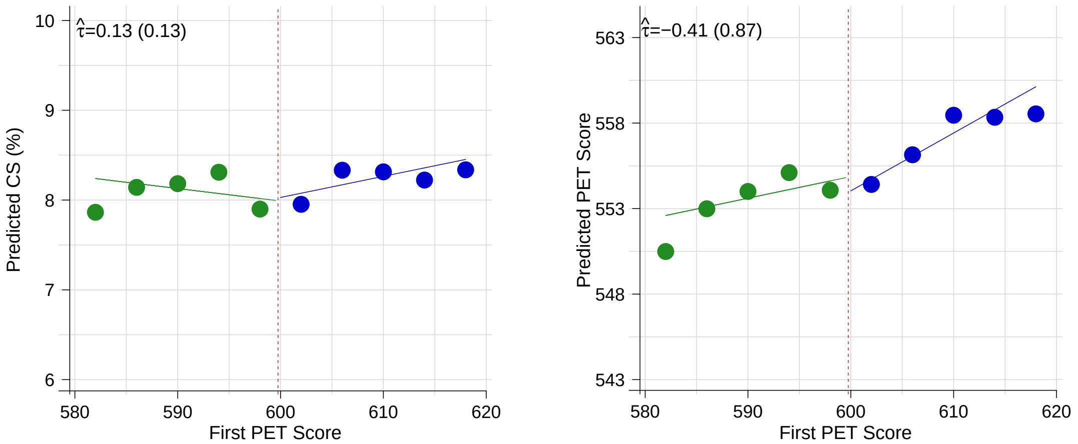
Notes: The figures present the testing regularities of Jews and Arabs in Israel. The sample includes all individuals in Israel who participated in their first PET during 1995-2008 (332,947 observations). The figure in the left panel (a) shows the distribution of ages at first test, by population group. The figure in the right panel (b) shows the distribution of number of tests taken by the individual by group.

Figure B.6: Descriptive Evidence on the Relationship between First PET Scores and Applications to Computer Sciences Within Three Years



Notes: The figures show the relationship between first PET scores and applications to Computer Sciences degrees in universities within three years after the test. The y-axis show the rate of applications to CS. The x-axis shows the first PET score (in bins of four points). The sample includes all individuals in Israel from cohorts 1979 and later, who participated in their first PET during 1995-2008 and apply to Computer Sciences (332,947 observations). The red dashed vertical line represents the score cutoff of 600.

Figure B.7: The "Impact" of Crossing 600 on Predicted Outcomes



(a) Predicted CS (b) Predicted Score

Notes: The figures show the estimated discontinuity in *predicted* outcomes at the score cutoff of 600 in the first PET. The y-axis show the predicted rate of CS applications (a) and the predicted PET score (b) based on all pre-determined outcomes showin in Table 1. The x-axis shows the first PET score (in bins of four points). The sample includes all individuals in Israel who participated in their first PET during 1995-2008 (43,911 effective observations). The red dashed vertical line represents the score cutoff of 600. The blue and green solid lines are based on the estimation of equation 1. The figures also report the estimates (and robust standard errors) for the coefficient of interest, τ, which represents the impact of crossing 600 on the outcome variable.

Appendix C Tables

Table C.1: Most Common Fields of Study in Universities in Israel

	N	Earnings (10,000 NIS)	Tech (%)	Q. Score (50-15)	T. Score (200-800)	Females (%)	Dad Educ. (years)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
CS	8,610	29.9	34.3	131.2	649.2	29.8	14.9
EE	7,349	22.4	45.5	133.0	652.4	14.5	15.0
Nursery	6,305	20.3	0.5	101.6	496.3	79.4	12.8
Management	5,995	18.0	16.0	113.9	565.4	58.8	13.6
Economics	12,313	17.0	11.9	123.7	611.8	44.9	14.2
Law	8,606	14.9	2.7	123.2	638.1	59.1	14.8
Politics	6,465	12.2	9.5	109.9	573.0	60.5	14.1
Social Work	5,473	11.9	1.8	108.3	555.5	91.8	13.6
Sociology	6,324	10.9	8.6	106.2	547.2	85.7	13.5
Psychology	9,476	10.9	8.2	119.5	620.6	79.1	14.7
Social Sciences	7,273	10.9	5.6	88.9	423.7	79.2	11.5
Humanities	6,816	10.2	8.2	102.6	515.9	70.5	13.2
Biology	8,334	9.9	15.2	120.4	614.0	72.4	14.7

Notes: The table shows information on the most common fields of study in universities in Israel (with at least 5,000 graduates in our data). The sample included all individuals in Israel who graduated with a university degree during the sample period. Column (1) shows the number of graduates in our sample. The other columns, (2)-(7), represent the mean values of the following outcomes among each field's graduates: Average annual earnings at the age of 28, rate of employment in the tech industry at the age of 28, the average PET score in the quantitative domain, the average PET total score, share of females, and the average years of education of the father.

Table C.2: The Impact of Crossing 600 on University Applications, Sample of All Tests

Outcome	Mean (1)	Estimate (2)
A. Within three years		
Computer Sciences (%)	5.99	1.106***
		(0.410)
Electrical Engineering (%)	4.63	0.956***
		(0.349)
Any (%)	51.86	0.71
		(0.803)
STEM (%)	28.83	1.973***
		(0.748)
Non STEM (%)	23.02	−1.264*
		(0.68)
Predicted income (10,000 NIS)	23.67	0.402**
		(0.164)
B. Ever		
Computer Sciences (%)	9.53	1.62***
		(0.499)
Electrical Engineering (%)	0.907**	
		(0.419)
Any (%)	71.90	−0.613
		(0.664)
STEM (%	41.40	2.038**
		(0.799)
Non STEM (%)	30.45	−2.344***
		(0.741)
Predicted income (10,000 NIS)	24.36	0.371***
		(0.14)
Number of observations		43,911

$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Notes: The table shows the estimated impact of crossing the score cutoff of 600 in the PET on university applications. Columns (1) and (2) show the baseline mean, and the estimates for τ in equation 1 (cluster-robust standard errors in parenthesis). The sample is based on all individuals in Israel who participated in a PET during 1995-2008, and it includes only individuals with first test scores within the window of 580 and 619.

Table C.3: The "Impact" of Crossing Non-round Score Cutoffs on CS Applications

Cutoff	580	584	588	592	596	600	604	608	612	616	620
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Within three years	−0.510	−0.069	0.016	−0.370	0.313	1.397***	0.081	−0.650	−0.510	−0.116	0.298
	(0.370)	(0.379)	(0.382)	(0.391)	(0.385)	(0.427)	(0.460)	(0.459)	(0.476)	(0.470)	(0.500)
Ever	−0.267	0.410	−0.130	−0.313	0.034	1.719***	0.582	−1.711***	−0.318	−0.283	0.946
	(0.528)	(0.549)	(0.550)	(0.562)	(0.555)	(0.590)	(0.628)	(0.625)	(0.629)	(0.625)	(0.662)

$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Notes: The table shows the results of the placebo exercise, estimating the impact of crossing non-round score cutoffs. The numbers are the estimate for τ in equation 1, where the outcome variable is mentioned on the leftmost column, and the cutoff used in the analysis is mentioned at the top row (robust standard errors in parenthesis). The sample is based on all individuals in Israel who participated in a PET during 1995-2008, and it includes only individuals with first test scores within the window of 20 points below and above the cutoff.

Table C.4: Falsification Tests, by Group

	Arabs		Jews, Age≤20		Jews, Age>20	
Outcome	Mean (1)	Estimate (2)	Mean (3)	Estimate (4)	Mean (5)	Estimate (6)
A. Individual						
Age	17.43	0.091	18.03	−0.017	22.30	−0.070*
		(0.115)		(0.033)		(0.041)
Female share (%)	50.04	0.066	59.44	−2.312*	51.70	−1.275
		(4.776)		(1.375)		(1.523)
Non-rel. school (%)	99.91	−0.071	78.20	−0.026	86.17	0.344
		(0.068)		(1.171)		(1.055)
Born in Israel (%)	97.74	−1.162	77.60	−1.720	90.28	0.144
		(1.614)		(1.168)		(0.909)
B. Parents						
Born in Israel (%)	93.95	0.764	37.12	0.604	44.29	1.102
		(2.317)		(1.358)		(1.516)
Income > 250K NIS (%)	21.14	0.764	44.13	−1.487	42.48	−0.452
		(4.014)		(1.395)		(1.511)
Post HS educ. (%)	31.89	6.248	56.60	−0.632	44.60	1.510
		(4.547)		(1.383)		(1.517)
C. Test						
Year	2004.32	−1.06***	2002.04	0.053	2005.01	−0.125*
		(0.335)		(0.100)		(0.069)
Month	7.45	−0.235	8.49	−0.116	5.97	0.012
		(0.333)		(0.086)		(0.106)
D. Predicted Values						
Predicted CS	8.09	0.157	11.17	0.174	4.21	0.069
		(0.512)		(0.200)		(0.079)
Predicted PET score	442.95	3.466	555.72	−0.630	562.59	0.973
		(3.545)		(1.011)		(1.060)
Number of observations	1,892		22,963		19,054	

$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Notes: The table shows the estimated discontinuous change in pre-determined outcomes at the score cutoff of 600 in the first PET, among different groups of test-takers. Columns (1), (3) and (5) show the baseline mean. Columns (2), (4), and (6) shows the estimates for τ in equation 1 (and their robust standard errors). The sample in columns (1) and (2) includes only Arab test-takers. The sample in columns (3) and (4) includes only Jewish test-takers, who were at age 20 or below when taking their first test. The sample in columns (5) and (6) includes only Jewish test-takers, who were at age 21 or above when taking their first test. The full sample is based on all individuals in Israel who participated in their first PET during 1995-2008, and it includes only individuals with first test scores within the window of 580 and 619.

Table C.5: Heterogeneous Impact of Crossing 600 on CS Applications, Sample of Younger Test-Takers (Ages 20 and below)

A. By gender	Men		Women		Difference	
	Mean	Estimate	Mean	Estimate	Baseline	Impact
	(1)	(2)	(3)	(4)	(5)	(6)
Three Years (%)	6.751	2.035*	2.644	1.533**	−4.107***	−0.502
		(1.083)		(0.640)	(0.784)	(1.258)
Ever (%)	19.12	3.256**	7.308	1.737*	−11.812***	−1.519
		(1.648)		(0.949)	(1.281)	(1.901)
Number of observations	9,676		13,590			
B. By SES	Low		High		Difference	
	Mean (1)	Estimate (2)	Mean (3)	Estimate (4)	Baseline (5)	Impact (6)
Three Years (%)	4.717	1.387*	3.371	2.751***	−1.346*	1.364
		(0.722)		(0.994)	(0.738)	(1.229)
Ever (%)	12.299	1.597	11.609	4.694***	−0.690	3.098
		(1.073)		(1.600)	(1.275)	(1.927)
Number of observations	9,833		13,433			

$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Notes: The table shows the estimated heterogeneous effects of crossing 600 in the first PET on university applications. It shows estimated effects on different groups of individuals, by their gender (A), and SES (B). The sample of high (low) SES includes all individuals for whom both parents have education beyond high school (13 years or more). Columns (1) and (3) show the baseline mean. Columns (2) and (4) show the estimates for the impact of crossing 600 on the outcome. Columns (5) and (6) report the difference in the baseline mean (5) and the effect (6) between the groups.

Table C.6: The Impact of Crossing 600 on Long Term Outcomes, Sample of Older Test-takers (21 and Above)

Outcome	Mean (1)	Est. Effect (2)
A. CS Degrees		
Enrollment (%)	1.25	−0.316
		(0.354)
Attainment (%)	0.54	−0.020
		(0.231)
Attainment, including college degrees (%)	2.45	0.085
		(0.490)
B. Labor market, at age 28		
Log Annual Earnings × 100	1109.98	−1.649
		(4.126)
Employment (%)	87.01	−1.55
		(1.019)
Self employment (%)	5.37	1.472**
		(0.697)
Employment in the tech industry (%)	10.32	0.539
		(0.957)
C. Labor market, 12 years after the test		
Earnings rank, conditional on age (%)	46.09	−1.132
		(0.977)
Employment (%)	97.07	−3.753**
		(1.747)
Self employment (%)	16.20	−0.692
		(1.289)
Employment in the tech industry (%)	8.18	0.897
		(0.977)
Number of observations		19,054

$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Notes: The table shows the estimated impact of crossing the score cutoff of 600 in the first PET on long-term outcomes. Columns (1) and (2) show the baseline mean and the estimates for τ in equation 1 (robust standard errors in parenthesis). The sample is based on all individuals in Israel who participated in their first PET during 1995-2008 when they were at age 21 or above, and it includes only individuals with first test scores within the window of 580 and 619.

Table C.7: The Impact of Crossing 600 on Long Term Outcomes, Sample of Arab Test-takers

Outcome	Mean (1)	Est. Effect (2)
A. CS Degrees		
Enrollment (%)	9.85	0.045
		(2.838)
Attainment (%)	4.70	−2.78
		(1.858)
Attainment, including college degrees (%)	5.15	−2.679
		(2.003)
B. Labor market, at age 28		
Log Annual Earnings × 100	1152.66	12.943
		(11.295)
Employment (%)	98.37	-12.185^{**}
		(4.987)
Self employment (%)	10.04	2.103
		(3.102)
Employment in the tech industry (%)	6.78	−1.246
		(2.597)
C. Labor market, 12 years after the test		
Earnings rank, conditional on age (%)	50.27	0.698
		(3.106)
Employment (%)	97.74	−6.642
		(5.038)
Self employment (%)	12.56	1.638
		(3.454)
Employment in the tech industry (%)	6.41	−1.523
		(2.569)
Number of observations		19,054

$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Notes: The table shows the estimated impact of crossing the score cutoff of 600 in the first PET on long-term outcomes. Columns (1) and (2) show the baseline mean and the estimates for τ in equation 1 (robust standard errors in parenthesis). The sample is based on all individuals in Israel who participated in their first PET during 1995-2008 when they were at age 21 or above, and it includes only individuals with first test scores within the window of 580 and 619.

Table C.8: The Impact of Crossing 600 on Additional Long Term Outcomes, Sample of Younger Test-takers (20 and Below)

Outcome	Mean (1)	Est. Effect (2)
A. Degrees, including college degrees		
STEM (%)	34.40	2.209
		(1.346)
Non STEM (%)	51.57	-2.650^{*}
		(1.400)
Any field (%)	85.98	-0.441
		(0.939)
B. Labor market, age 28		
Annual earnings (10,000 NIS)	14.68	1.622^{*}
		(0.910)
Annual earnings, employed=1 (10,000 NIS)	17.24	1.830^{*}
		(1.049)
Annual salaried earnings (10,000 NIS)	14.26	1.704^{*}
		(0.900)
C. Labor market, age 29		
Log earnings × 100	1149.45	6.300^{*}
		(3.789)
D. Marriage		
Married, age 27 (%)	30.33	-0.714
		(1.286)
Married, 2019 (%)	68.57	-0.505
		(1.308)
Number of observations		22,963

$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Notes: The table shows the estimated impact of crossing the score cutoff of 600 in the first PET on long-term outcomes. Columns (1) and (2) show the baseline mean and the estimates for τ in equation 1 (robust standard errors in parenthesis). The sample is based on all individuals in Israel who participated in their first PET during 1995-2008 when they were at age 20 or below, and it includes only individuals with first test scores within the window of 580 and 619.

Table C.9: The "Impact" of Crossing Non-round Score Cutoffs on Long-term Outcomes, Sample of Younger Test-takers (20 and below)

Cutoff	580	584	588	592	596	600	604	608	612	616	620
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
CS enrollment	0.137	0.31	−0.546	−0.439	0.632	1.809***	−0.263	−2.264***	1.303*	−0.127	0.956
	(0.583)	(0.619)	(0.613)	(0.629)	(0.62)	(0.686)	(0.731)	(0.701)	(0.687)	(0.731)	(0.788)
CS degree	−0.548	0.733	−0.12	−0.481	0.457	0.972*	−0.238	−1.411**	0.561	0.214	1.322**
	(0.455)	(0.492)	(0.506)	(0.514)	(0.52)	(0.57)	(0.606)	(0.593)	(0.579)	(0.596)	(0.65)
Earnings rank	0.861	−0.235	−1.165	−1.345	0.126	1.731*	−0.025	−1.636*	−1.107	−1.559	1.686*
	(0.902)	(0.922)	(0.91)	(0.914)	(0.925)	(0.945)	(0.959)	(0.942)	(0.97)	(0.968)	(1.016)

$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Notes: The table shows the results of the placebo exercise, estimating the impact of crossing non-round score cutoffs. The numbers are the estimate for τ in equation 1, where the outcome variable is mentioned on the leftmost column, and the cutoff used in the analysis is mentioned at the top row (robust standard errors in parenthesis). The sample is based on all individuals in Israel who participated in a PET during 1995-2008 when they were ages 20 and below, and it includes only individuals with first test scores within the window of 20 points below and above the cutoff.

Table C.10: The Impact of Crossing 600 on Long Term Outcomes, Heterogeneity by Gender and SES, Sample of Younger Test-takers (20 and Below)

A. By Gender	Men		Women		Difference	
	Mean (1)	Estimate (2)	Mean (3)	Estimate (4)	Baseline (5)	Impact (6)
CS degree (%)	6.286	1.318	2.676	0.259	-3.61^{***}	-1.060
		(1.019)		(0.548)	(0.763)	(1.157)
Earnings rank (%)	48.497	3.494^{**}	45.6	0.945	-2.897^{**}	-2.548
		(1.494)		(1.085)	(1.293)	(1.846)
Number of observations	9,676		13,590			
B. By SES	Low		High		Difference	
	Mean (1)	Estimate (2)	Mean (3)	Estimate (4)	Baseline (5)	Impact (6)
CS degree (%)	3.92	0.740	4.653	0.742	0.732	0.003
		(0.629)		(0.989)	(0.770)	(1.172)
Earnings rank (%)	46.973	1.989^{*}	46.445	2.078	-0.528	0.090
		(1.069)		(1.594)	(1.338)	(1.920)
Number of observations	9,833		13,433			

$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Notes: The table shows the estimated heterogeneous effects of crossing 600 in the first PET on university applications. It shows estimated effects on different groups of individuals, by their gender (A), and SES (B). The sample of high (low) SES includes all individuals for whom both parents have education beyond high school (13 years or more). Columns (1) and (3) show the baseline mean. Columns (2) and (4) show the estimates for the impact of crossing 600 on the outcome. Columns (5) and (6) report the difference in the baseline mean (5) and the effect (6) between the groups.