The Long-Term Effects of Gifted Programs*

Niv Brosh, Oren Danieli, David Maagan, Ran Shorrer, Noam Zussman†

September 25, 2026

Preliminary and incomplete — do not distribute

Abstract

We evaluate three selective enrichment programs in Israel designed for students at the top of the ability distribution. Using a regression discontinuity design based on admission-test cutoffs, we estimate the causal effects of admission to full-time, weekly pull-out, and afternoon programs. Our analysis links historical admission-test records, beginning in the late 1980s, to administrative data on education, labor income, family formation, startup employment, as well as external data on patents and academic publications. We find that effects on short-term academic achievement range from zero to negative, whereas effects on longer-run outcomes are positive. Admission increases higher-education attainment, particularly in STEM fields, as well as measures of innovation, including patents, startup employment, and scientific publications. We additionally find positive effects on marriage rates. The effects are often larger for women. Finally, we find no evidence of negative effects on the classmates left behind. Our findings support arguments in favor of gifted education: they benefit participants in a variety of ways, as well as the greater society by fostering innovation.

*We are grateful to Analia Schlosser, seminar participants at Tel Aviv University, the NBER economics of talent meeting, LMU, Northwestern, and BSE summer forum for helpful comments and suggestions. We thank Menachem Nadler, Director of the Ministry of Education's Department for Gifted and Outstanding Students, former department staff member Gilmor Keshet, and Rinat Yitzhak of the Szold Institute for their assistance in making the program-admission test scores available and for their explanations of the programs. We thank the staff of the Israel Central Bureau of Statistics for preparing the research data and research assistants Arielle Lubick and Idan Yehud. We gratefully acknowledge financial support from the Spencer Foundation. Any views expressed in this paper are those of the authors and do not necessarily reflect those of the Bank of Israel.

†Danieli (corresponding author): Tel Aviv University and CEPR, orendanieli@tauex.tau.ac.il; Brosh: Hebrew University; Maagan: Israel Central Bureau of Statistics; Shorrer: Penn State University; Zussman: Bank of Israel.

Specialized programs for gifted students have been widespread across the world for decades, yet the debate over whether they should exist remains as fierce as ever. Critics contend that gifted programs use public funding to disproportionately benefit already-advantaged students (Borland, 2003), lack racial and socioeconomic diversity (Card and Giuliano, 2016a; Thompson, 2024), and negatively affect peers who remain in regular classrooms (Burris and Garrity, 2008). They also cite research showing that these programs may not benefit participants (Bui et al., 2014; Boisjoly et al., 2006). On the other hand, proponents (e.g., Rutigliano and Quarshie, 2021) argue that gifted programs cultivate talent that benefits society through innovation and scientific discovery (Subotnik et al., 2011), support gifted students who may lag behind the general population in socio-emotional development (Silverman, 2017), and promote equal opportunity by helping high-ability students from underrepresented backgrounds realize their potential, thus avoiding "lost Einsteins" and "lost Marie Curies" (Bell et al., 2019; Hoisl et al., 2023; Olszewski-Kubilius et al., 2004; Plucker et al., 2010; Card et al., 2026). They also contend that the evidence on limited program effectiveness is due to focus on short-run test scores and not on the programs' stated objectives.

A high-profile illustration of these tensions is the repeated reversals in gifted-education policy in New York City. After expansion under Mayor Bloomberg, Mayor de Blasio moved to phase out gifted programs. His successor, Mayor Adams, reversed course and expanded gifted programs citywide. However, the current mayor, Mamdani, campaigned on ending gifted programs in early grades once more.[1] Despite the intensity of this debate, empirical evidence on gifted education, especially in the long run, remains limited for two main reasons. First, assignment to gifted programs is highly non-random. Second, evaluating long-term effects requires matching historical student records from longstanding programs with later life outcomes.

We overcome these challenges by collecting historical admissions data from three national gifted programs in Israel and merging it with multiple data sources covering a broad range of outcomes, and implementing a regression discontinuity design (RDD). Our results support the main arguments in favor of gifted programs and reject some of the key arguments against them. Consistent with previous studies, we find null or negative effects on short-

[1]For coverage by *The New York Times*, see ""Schools to Add More Programs for the Gifted"; "De Blasio to Phase Out N.Y.C. Gifted and Talented Program"; "Adams Commits, With Few Details, to Keeping Gifted Program in Schools"; and "Mamdani Says He Would Phase Out N.Y.C. Gifted Program for Early Grades."

term academic achievement. However, focusing on short-term academic outcomes misses substantial broader and longer-run benefits: Admission increases higher-education attainment, particularly in STEM fields. It also increases patenting, startup employment, and academic publications, speaking directly to proponents' argument that the broader society benefits from programs' innovation externalities. In addition, admission increases marriage rates—a dimension in which the gifted population lags behind the general population. The effects are often larger for women, consistent with the view that gifted programs may help prevent "lost Marie Curies." Finally, we find no evidence that gifted programs harm the classmates left behind in regular classrooms.

The three programs we study are centrally designed and monitored by the Ministry of Education (MOE), but differ in the population they target, their format, and their intensity. The first two programs target approximately the top 1–3 percent of the ability distribution. The most intensive program is the *full-time program*. This program brings together gifted children from across the district to form a new class where students study the standard curriculum in addition to an enrichment program. This program is offered only in a small number of school catchment areas. In most catchment areas, the gifted program takes the form of a weekly *pull-out* day. Participants in this program are pulled out from their regular classes one day a week and attend enrichment classes in a specialized center. They continue to study the standard curriculum in their original classroom in their neighborhood school. The third program we study, *the afternoon program*, targets the next 3–5 percent in the distribution of ability. It offers a similar enrichment curriculum but at a lower intensity, and takes place after regular school hours.

To capture the broad impacts of these programs, we construct a comprehensive dataset that covers the admission process, a wide array of outcomes over a long horizon, and a rich set of family and background characteristics. We collect records from the national admission exam, held in the second grade, beginning in 1988. We link them to administrative records from several government agencies. These records cover individuals' educational trajectories from elementary-school test scores through graduate degrees, labor-market outcomes, demographic outcomes, and links to parents, spouses, children, and classmates. We augment these data with Google Scholar publication records and United States patent records available through PatentsView. This combination provides a broad view of the candidates'

backgrounds and long-run outcomes.

Our identification strategy relies on strict cutoff-based admissions. For each program, we aggregate results from all locations across all years it was offered using a pooled estimator (Cattaneo et al., 2016). We estimate intention-to-treat (ITT) effects because for the early cohorts, which permit long-run follow-up, we only observe admission decisions but not actual participation. Data from recent cohorts indicates that the effect of admission on participation is between 46 and 85 percentage points (p.p). The multiple-choice admission exam is graded by an external NGO that also releases the admission decisions. The admission score itself remains confidential: students, parents, and teachers learn the admission decision but not the score.

We begin by analyzing short- and medium-run educational outcomes. Consistent with much of the previous literature, we find no positive effect on standardized test scores. The effects of admission to the full-time and the afternoon programs on short-run test scores are close to zero. Admission to the pull-out program reduces fifth-grade test scores, particularly for females. This is plausibly because participants miss a full day of regular instruction each week. This gap disappears by the eighth grade, and we find no persistent losses in high-school outcomes.

The null results in the short and medium run stand in sharp contrast with the longer-run results. Admission increases the probability of completing undergraduate and graduate degrees, with particularly strong effects on STEM education. The more intense full-time program generally yields the largest effects. Direct effects on labor income is imprecisely estimated. This is not surprising in our context, as military service and prolonged higher education delay labor-market entry. A surrogate exercise that predicts later labor income based on educational attainment (field and institution) displays positive income effects (Athey et al., 2026).

A central contribution of our paper is the finding that gifted programs substantially increase the innovative activity of participants. We measure innovative activity along three dimensions. First, we find large effects on patenting among students admitted to the full-time program. Admission increases the probability of becoming an inventor by 5 p.p., from a baseline of 6 p.p. Second, we find an effect of 9 p.p. of admission to the full-time program on employment in startups. Third, we measure scientific activity using whether an individual has

a Google Scholar profile and the number of citations received. We cannot detect an overall statistically significant effect on our measures of scientific activity. However, in the pull-out program, where the larger sample provides greater statistical power, admission increases both the probability of having a Google Scholar profile and the number of citations among women, with no corresponding effect on men. These findings suggest that gifted programs have impacts that are beyond the participating students, who are already advantaged. They also create valuable externalities for society.

Admission to these elementary-school programs also increases marriage rates decades later. However, admitted students still marry at lower rates than the general population: by age 30, 40–51 percent of students across the three programs are married, compared with 57 percent of the general population. This is notable because marriage is not an outcome typically affected by educational interventions delivered at such young ages.

We investigate two broad mechanisms for the effect on marriage rates. First, the programs may alter participants' social networks by changing their educational path. Second, gifted programs aspire to develop socio-emotional skills, which predict higher marriage rates (Lundberg, 2012; Heckman et al., 2006). The evidence is suggestive of the first mechanism for boys in the full-time program and of the second for girls.

We next explore heterogeneity in the effects of these programs in the context of the equity debate. In the older cohorts that our long-term estimates rely on, boys and students from high socioeconomic backgrounds are disproportionately represented among admitted students. However, girls often benefit more from admission, particularly in the pull-out program. Thus, admitting a more diverse cohort can potentially improve the programs' average impact.[2]

Finally, we examine effects on students who were classmates of applicants before the admission decision. Critics worry that removing high-ability students may harm those left behind (Burris and Garrity, 2008). We find no evidence of negative average effects on classmates' test scores, high-school GPA and credits, or graduate-degree outcomes. Furthermore, our confidence intervals rule out meaningful effects. Moreover, among classmates from lower socioeconomic backgrounds, we find positive effects on test scores of classmates of students admitted to the afternoon program.

[2]In later cohorts, diversity in admissions substantially increased following an affirmative action policy, which we discuss in Section 1.2.

Relation to the Literature This paper makes five key contributions to the literature on gifted education. First, our unique setting allows us to study long-term impacts, which more closely reflect the programs' stated objectives. Previous studies of selective gifted and talented programs have focused primarily on short-term outcomes, often within a single U.S. school district (Abdulkadiroglu et al., 2014; Bui et al., 2014; Card and Giuliano, 2014; Card and Giuliano, 2016b; Dobbie and Fryer, 2014; Redding and Grissom, 2021; Cleveland, 2023; Chin and Kocks, 2025). Yet improving short-term outcomes such as test scores is not a primary objective for most gifted programs (Renzulli, 2021), which may explain why this literature generally finds small or null effects on test scores.[3]

Second, we provide credible causal evidence on these long-term effects using an RDD. Most papers that study long-term outcomes do not observe admission scores, which are necessary for such design. Several papers descriptively characterize the long-term outcomes of children who participated in gifted programs (Gross, 2006; Wai et al., 2010; Helal et al., 2026). Others are using matching methods to address the absence of a natural experiment (Freeman, 2006; Bernstein et al., 2021). A contemporaneous paper by Goldstein et al. (2026) uses a matching design to study elite high-school classes in Israel. It also implements a difference-in-differences design that compares across localities with "foundational" gifted programs to localities where they are not offered, yet may have alternative programs for high-performing students. By contrast, in the early cohorts of our sample, which we use to study long-term effects, students who were not admitted had no access to alternative gifted programs, reducing concerns that the control group received close substitutes (Kline and Walters, 2016).

Third, we study outcomes that speak not only to the private benefits of gifted programs for participants, but also to their broader social consequences and externalities. Cohodes (2020) finds that an accelerated curriculum program in Boston improves college attainment. Booij et al. (2016) find that a pull-out program in the Netherlands increases the choice of fields with higher returns. Goldstein et al. (2026) also find an impact on advanced degrees, field of study, and on the spouse "quality" (but not on marriage rates). These findings critically inform private decisions of students and parents who consider participation in these programs. We contribute to this literature by expanding the scope to innovative activity,

[3]Davis et al. (2010) provide evidence that a gifted program offered in an American public school district increases the probability that students from higher-income families remain in the public system.

which is a key motivation for public investment in such programs. We also study effects on non-gifted classmates. Previous papers have used quasi-random student assignment to measure the effect of having a high-achieving classmate (Card and Giuliano, 2016a; Moreira, 2016; Balestra et al., 2023). We estimate the direct effect of removing gifted students on the students left behind.

Fourth, we study programs targeted at the extreme upper tail of the ability distribution. This population is particularly relevant for studying innovation. For example, students just below the admission cutoff in the full-time program have a 4% chance of filing a patent application, far above the population average. Previous work by Cohodes (2020) and Goldstein et al. (2026) study programs that target a broader population.[4] Fifth, by studying three different programs, we provide evidence relevant to dilemmas on how gifted programs should be implemented (Rogers, 2007; Steenbergen-Hu et al., 2016; Mammadov and Hertzog, 2024). Specifically, the larger effects of the full-time program suggest that program intensity may matter. Our findings stand in contrast to the expert opinion in an official MOE report, which determined that a weekly pull-out day is most appropriate (Noy and Vorgan, 2026).

1 Institutional Background

1.1 Program Descriptions

We study three national programs for gifted and outstanding students: the *full-time* program, the *pull-out* program, and the *afternoon* program. Since we were not able to retain historical admission records for Arabic-language admissions exams, we focus on programs in the Hebrew-speaking systems (public and public religious). Together, these systems covered approximately 70% of the country's students.[5]

All three programs we study are centrally designed and monitored by the MOE. They are offered during elementary school, primarily in grades 3–6. The programs provide enrichment in subjects that do not overlap with the regular school curriculum. In particular, the programs are not designed as acceleration programs. In addition to the enrichment curriculum, all three programs expose students to high-ability peers. Admission to the programs also provides

[4] See footnote 2 in Cohodes (2020) for detailed comparison of targeted populations across different studies.
[5] Until recently, there were no gifted programs for the ultra-orthodox population.

students and their environment with a salient signal on the students' ability. The differences between the three programs are summarized in Table 1.

The full-time gifted program is offered only in a small number of affluent municipalities. Participants in this program transfer to a designated elementary school, which is otherwise a regular school and is typically located outside their neighborhood. They study in a full-time gifted class that combines the standard curriculum with enrichment curriculum throughout the school week. Although this program is not designed as an acceleration program, it may allow participants to study the standard school curriculum more quickly and deeply, which is not possible in the other two programs. Participation in this program is free of charge.

The pull-out program is the largest gifted program, and it is available in most municipalities. This program consists of a weekly pull-out day in which students attend a gifted center rather than their regular school. Students therefore miss a full day of regular instruction each week, which they are expected to make up independently. Each pull-out day consists of three enrichment sessions that cover topics outside the standard school curriculum (see Appendix Fig. A.1 for an illustrative schedule). The program requires co-payments of approximately \$50 per month (slightly less than 1% of the average household income). Some municipalities provide bussing.

The afternoon program is offered in many but not all districts. It takes place after regular school hours, often in the same gifted centers. While the other two programs are labeled as gifted programs, participants in this program are labeled outstanding, as they meet a lower admission bar (details on the admission process are provided below). Participants in this program receive one or two weekly enrichment sessions (compared to three sessions in the pull-out program). Transportation is not provided, and families pay tuition.

1.2 The Admission Process

Admission is determined through a two-stage national process. First, all second-grade students take a preliminary screening examination (in a few localities, screening takes place in third grade). Approximately 10–15 percent are referred to the main admission examination based on their performance in the exam or based on teacher or parents' recommendation.

The main admission examination is a multiple-choice, paper-based intelligence test. It contains five equally weighted sections that cover several forms of mathematical and logical

Table 1: The Three Programs

	Full-time	**Pull-out**	**Afternoon**
Timing	All school hours	One day per week	One day per week
Misses regular school	No	Yes	No
Curriculum	Enrichment and standard curriculum	Three enrichment sessions	One or two enrichment sessions
Target population	Top 1–3%	Top 1–3%	Next 3–5%
Availability	A few large municipalities	Most municipalities	Most municipalities
Location	Designated elementary school	Gifted center	Gifted center
Transportation	Provided in some municipalities	Provided in some municipalities	Not provided
Price	Free	Parental co-payment	Parental co-payment

Notes: The table summarizes the institutional features described by the MOE. The target-population rows are approximate: actual cutoffs vary by center, cohort, gender, and local capacity. With a few exceptions in recent cohorts, municipalities offered either a full-time or pull-out gifted program, not both.

reasoning, language, and general knowledge. The exam is designed, administered, and graded by an NGO that was retained by the MOE for this purpose. The NGO also issues the admission decision based on the student's score relative to the relevant cutoff. Students, parents, and teachers observe the final decision but not the exact score. The centerpiece of our admissions data is the NGO's internal data, which includes the exact admission score as well as the admission decision.

Admission-exam scores strongly predict later-life outcomes. For example, Appendix Fig. A.2 shows a scatter plot of the raw test score and the share of students who obtained a PhD by age 34. Below a score of 60, the share is approximately 1%. This share then increases approximately linearly with the test score, reaching 8% for the highest levels. This illustrates that this second-grade exam can predict significant life choices and achievements decades later.

In the full-time and pull-out programs, the cutoffs are chosen to admit approximately the top 1–3 percent of students in the catchment area. In a catchment area where an afternoon program is available, the cutoff is set to accept approximately the next 3–5 percent. Hence, students who score slightly below the full-time or pull-out program cutoff will be admitted

to the afternoon programs in locations where such program is available (see Section 3.1 for details). In practice, the cutoff varies across gifted centers and cohorts because admission is constrained by local capacity. In 2014, the MOE introduced an affirmative action policy that sets differential admission thresholds based on the socio-economic index of candidates' elementary school.

Admission thresholds may also vary by gender. In 2001, an official national affirmative action policy was introduced in the full-time and pull-out programs, substantially increasing the share of female participants. Before this policy change, lower thresholds were set in some centers ad hoc to promote gender balance. For this reason, when we calculate admission thresholds, we do so separately for each gender and location (see Section 3 for details).

2 Data

Our research uses a novel dataset that we assembled from administrative and external sources. We collect the results of the national admission examination beginning in the 1978 cohort. The records include the student's name and national identification number, the associated gifted center (catchment area), gender, examination score, and admission decision. Older paper records were digitized and manually validated at the Israeli Central Bureau of Statistics (CBS). The match to the population registry is nearly perfect, except for corrupted records in 1994.

We merge the admission records with the population registry using national identification numbers. The registry provides date of birth, marital-status histories, and links to parents, spouses, siblings, and children. These links allow us to study family background, family formation, and outcomes of relatives.

We obtain primary- and secondary-school records from the MOE. These data include annual school and homeroom assignments, low-stakes national examinations in grades 5 and 8, and high-school matriculation examinations with information on subject and level of study. The annual student records allow us to identify applicants' second-grade homeroom classmates and examine spillovers on students who shared the same homeroom before admission decisions were made.[6]

[6]In Israel, students normally study all school subjects with their homeroom classmates. The composition of the homeroom is largely fixed throughout elementary school (i.e., in grades 1 to 6).

Higher-education records from the Council for Higher Education report enrollment by institution, field of study, and year, as well as completed degrees, completion year, and grade-point averages. We use these records to construct indicators for bachelor's, master's, and doctoral education, including STEM fields.[7] Data from the National Institute for Testing and Evaluation contain scores on the University Psychometric Entrance Test (UPET). This exam evaluates abilities in verbal reasoning, quantitative reasoning, and English, and its outcome is a key component in college admissions. We observe these scores for applicants and their spouses.

We measure labor-market outcomes using tax records. These data report employment, labor income (wage and self-employment), and three-digit industry codes. We use industry to identify employment in the technology sector and in startups.[8]

Finally, we use United States Patent and Trademark Office records, accessed through PatentsView, to identify inventors, patent applications or grants, and forward citations. We use Google Scholar records to measure academic publications and citations.

Outcome availability differs across cohorts. While our key outcome variables are available for all cohorts, certain outcomes (e.g., national fifth grade exams), which have not been collected historically, are only available for later years. Moreover, individuals must be old enough to reach the relevant outcome age. For age-specific outcomes, we restrict the sample to individuals who are sufficiently old to determine the outcome based on the latest data we obtained (e.g., to have an outcome on "married by age 30," the individual must be at least 30 in our data).

Data on participation in the pull-out and afternoon programs is only available starting in 2012. For the full-time program (which changes the student's homeroom assignment) it is available starting in 1996. In light of this limitation and our focus on long-run effects, we only measure the effect of admission (analogous to ITT) and not participation.

Table 2: Summary Statistics for Older Cohorts

Outcome	Tested, admitted			Tested, not admitted	Not tested
	Full-time	Pull-out	Afternoon		
High-school GPA (standardized)	1.292	1.068	1.024	0.644	0.089
	(0.436)	(0.576)	(0.579)	(0.716)	(0.932)
	[635]	[4,270]	[2,828]	[28,173]	[511,919]
BA by age 28	0.706	0.532	0.549	0.435	0.247
	(0.456)	(0.499)	(0.498)	(0.496)	(0.431)
	[694]	[4,856]	[3,233]	[34,457]	[806,761]
MA by age 32	0.341	0.251	0.235	0.168	0.085
	(0.474)	(0.434)	(0.424)	(0.374)	(0.280)
	[693]	[4,851]	[3,229]	[34,428]	[806,725]
PhD by age 34	0.100	0.061	0.066	0.027	0.009
	(0.300)	(0.239)	(0.248)	(0.163)	(0.096)
	[693]	[4,849]	[3,228]	[34,409]	[806,711]
Married by age 30	0.401	0.512	0.470	0.540	0.569
	(0.490)	(0.500)	(0.499)	(0.498)	(0.495)
	[683]	[4,749]	[3,171]	[33,772]	[773,720]

Notes: Each cell reports the mean, with the standard deviation in parentheses and the number of observations in square brackets. The table uses the 1978–1988 birth cohorts, so that the listed long-run outcomes are observed. The first three columns contain tested students admitted to the named program. "Tested, not admitted" includes students who took the main admission examination but were not assigned to any of the three programs. "Not tested" includes other students in the same birth cohorts. The comparisons are descriptive and should not be interpreted causally.

2.1 Summary Statistics

Table 2 presents descriptive statistics. To preserve a comparable sample throughout the table, we restrict attention to the 1978–1988 birth cohorts, who are old enough to observe the main long-run outcomes. The first three columns present the averages in the full-time, pull-out, and afternoon programs, separately. The next column shows students who passed the first screening stage and took the main admission exam, but were not admitted. The last column presents students who did not take the main admission exam.

Admitted students are positively selected. For example, students admitted to the full-time program have a high-school GPA of 1.3 standard deviations above the population average and higher rates of bachelor's, master's, and doctoral education than tested students who were not admitted. Students admitted to the full time program have a 10% chance of obtaining a PhD by age 34, compared to less than 1% in the population that was not tested. These

[7]The data includes degrees from Israeli universities and self-reported foreign degrees.

[8]The CBS technology-sector definition includes pharmaceuticals (industry 21); computers, electronic and optical products (26); aircraft, spacecraft, and related equipment (303); telecommunications (61); computer programming, consultancy, and related services (62); data processing, hosting, related activities, and web portals (631); research-and-development centers (720); and research and development in engineering and the natural sciences (721), under the 2011 industry classification.

Table 3: Representation by Gender and Family Background

Outcome	Tested, admitted			Tested, not admitted
	Full-time	Pull-out	Afternoon	
Female	0.316	0.344	0.300	0.420
	(0.465)	(0.475)	(0.458)	(0.494)
	[702]	[4,911]	[3,266]	[34,824]
Mother has MA	0.177	0.134	0.140	0.081
	(0.382)	(0.341)	(0.347)	(0.273)
	[702]	[4,911]	[3,266]	[34,824]
Mother has PhD	0.053	0.025	0.028	0.011
	(0.224)	(0.156)	(0.165)	(0.106)
	[702]	[4,911]	[3,266]	[34,824]
Father has MA	0.160	0.127	0.129	0.070
	(0.366)	(0.333)	(0.335)	(0.256)
	[702]	[4,911]	[3,266]	[34,824]
Father has PhD	0.047	0.034	0.036	0.018
	(0.212)	(0.182)	(0.185)	(0.134)
	[702]	[4,911]	[3,266]	[34,824]
Parental labor-income percentile (age 8)	0.636	0.542	0.568	0.485
	(0.280)	(0.273)	(0.277)	(0.270)
	[702]	[4,911]	[3,266]	[34,824]

Notes: Each cell reports the mean, with the standard deviation in parentheses and the number of observations in square brackets. The sample uses the 1978–1988 birth cohorts. The first three columns contain tested students admitted to the named program. Parental education is measured in the administrative education records, and parental labor-income percentile is measured at age 8, before the applicant's admission examination. The table documents representation and is not a causal comparison.

comparisons are descriptive and combine selection on the admission score with any causal effect of the programs.

The table also illustrates why family formation is a distinct margin in this population. Despite their high educational attainment, students admitted to the programs are less likely to be married by age 30 than students who were tested but not admitted. This pattern motivates our analysis of whether the programs affect marriage rates.

Table 3 documents two dimensions of unequal representation for the same cohorts. The table presents gender and parental characteristics for students in the three programs, as well as students who were tested but not admitted.[9] The female share is only 30–34 percent of admitted students in the older cohorts. Admitted students also come from more advantaged families: their parents have substantially higher graduate-education rates and higher income percentiles than tested students who were not admitted. The imbalance is especially pronounced in the full-time program.

[9]Parental characteristics are not available for students who were not tested.

3 Methodology

We estimate the causal effect of admission using an RDD based on national admission-test scores. We focus on the effect of admission, as data on participation is not available for early cohorts, as discussed in Section 2. In Section 3.1, we show that admission has a large effect on participation among younger cohorts across all programs.

All three programs use a cutoff rule. Although the programs aim to identify a fixed share of the national ability distribution, local capacity implies that the threshold is determined at the catchment area level and varies across cohorts. Gender-specific affirmative-action policies also generate different thresholds for boys and girls in some cohorts. This yields variation in thresholds at the program, catchment area, year, and gender level.

The administrative records do not report these thresholds directly. Therefore, for each cell (program type, catchment area, year, and gender), we infer the admission threshold by choosing a score that minimizes the number of recorded admission decisions inconsistent with score-based admission.[10] Since we have multiple cutoffs, we follow Cattaneo et al. (2016) and pool observations after centering each score at its inferred cutoff.

For outcome Y_i and centered admission score R_i (raw admission score minus the cutoff), we estimate local specifications of the form

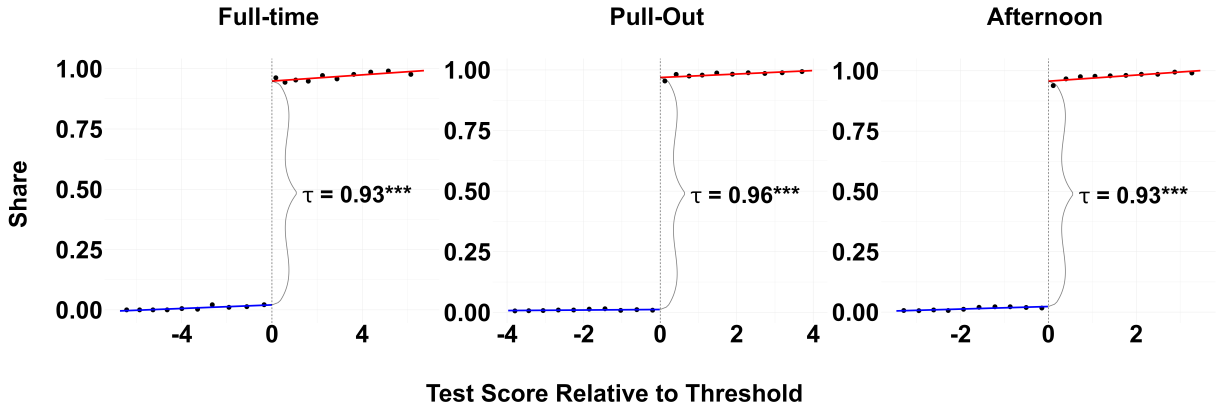
$$Y_i = \alpha + \tau \mathbf{1}\{R_i \geq 0\} + f_-(R_i)\mathbf{1}\{R_i < 0\} + f_+(R_i)\mathbf{1}\{R_i \geq 0\} + X_i'\gamma + \varepsilon_i, \tag{1}$$

where τ is the effect of crossing the admission cutoff and X_i is a vector of control variables including gender, indicator for religious school, parental MA status, and the raw test score (not normalized). Standard errors are clustered at the gifted-center-by-year level.

Figure 1 shows the first stage for recorded admission. Crossing the cutoff produces a large discontinuity in assignment for all three programs. The first stage is not perfectly sharp because a small number of recorded decisions deviate from the cutoff rule (likely due to data corruption or classification errors). A McCrary (2008) test finds no evidence of manipulation of the running variable around the cutoff. This is expected, as the exam is graded externally and the score and the admission thresholds are confidential. We also examine balance in predetermined covariates (negative controls, Danieli et al., 2024). Table 4 reports estimates for gender, family composition, maternal background, parental labor income, and school

[10] Since this definition usually yields a range of scores, we set the threshold in the middle of this range.

Figure 1: Admission at the Program Cutoffs



Notes: The figure plots the probability of recorded admission against admission-test score points relative to the inferred cutoff, separately for the full-time, pull-out, and afternoon programs. The cutoff in each cell defined by gifted center, year, and gender is chosen to minimize assignment misclassification and is centered at zero before pooling. Dots report binned means and the fitted lines summarize the relationship on each side of the cutoff. The discontinuity is the first stage for admission, not participation.

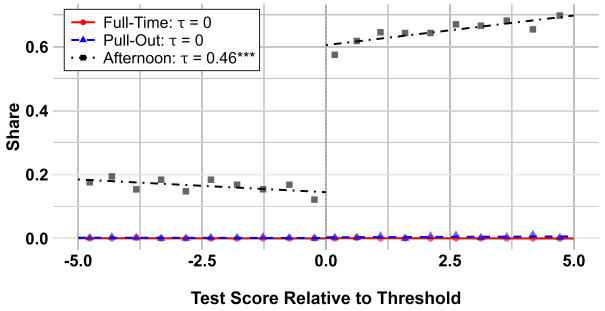
Table 4: Falsification Tests Using Predetermined Outcomes

	Full-time			Pull-out			Afternoon		
Outcome	Mean	(1)	(2)	Mean	(3)	(4)	Mean	(5)	(6)
		0.00	0.00		-0.01	-0.00		-0.01	-0.01
		(0.03)	(0.03)		(0.01)	(0.01)		(0.01)	(0.01)
Female	0.42	[5,190]	[5,443]	0.44	[44,488]	[44,000]	0.42	[28,703]	[28,637]
		-0.13	-0.11		0.03	0.01		-0.03	0.00
		(0.05)	(0.05)		(0.03)	(0.02)		(0.03)	(0.02)
Siblings born before exam	1.53	[4,427]	[4,454]	2.06	[38,984]	[44,904]	1.94	[43,988]	[44,232]
		0.03	0.03		0.00	0.00		0.03	0.02
		(0.02)	(0.02)		(0.01)	(0.01)		(0.01)	(0.01)
Mother born abroad	0.21	[6,430]	[6,722]	0.24	[46,313]	[44,036]	0.26	[35,324]	[35,467]
		-0.05	-0.04		-0.00	-0.00		-0.00	-0.00
		(0.03)	(0.03)		(0.01)	(0.01)		(0.01)	(0.01)
Mother has MA at exam	0.23	[4,223]	[4,302]	0.19	[40,047]	[42,250]	0.16	[43,047]	[38,323]
		-0.00	-0.00		0.00	0.00		-0.00	-0.00
		(0.02)	(0.02)		(0.01)	(0.01)		(0.01)	(0.01)
Parental labor-income percentile (age 8)	0.59	[5,416]	[5,430]	0.54	[37,312]	[35,309]	0.52	[44,768]	[41,643]
		-0.01	-0.01		0.00	0.00		0.01	0.01
		(0.03)	(0.03)		(0.01)	(0.01)		(0.02)	(0.01)
School SES mean	0.38	[6,537]	[7,335]	0.12	[46,671]	[48,130]	0.09	[46,874]	[38,551]
Controls		No	Yes		No	Yes		No	Yes

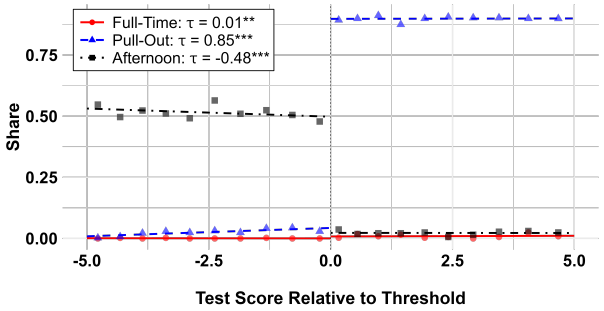
Notes: Each coefficient is a pooled RDD estimate for a characteristic determined before the admission examination and therefore not causally affected by admission. Standard errors are in parentheses and the number of observations is in square brackets. Columns labeled "Yes" include the study's predetermined controls. Standard errors are clustered at the gifted-center-by-year level. The absence of a systematic pattern is more informative than any single coefficient.

socioeconomic status. Out of 18 tests, only one (siblings born before exam in the full-time program) is significant at the 5% level. The overall pattern is consistent with local comparability around the inferred cutoffs.

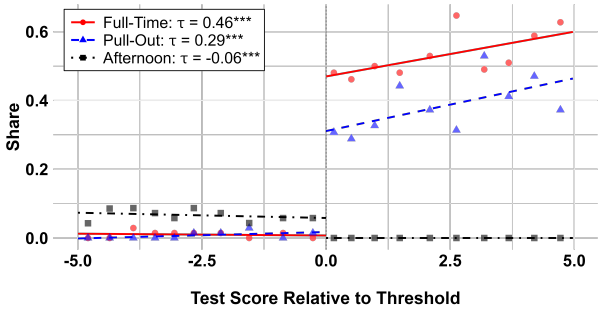
Figure 2: Effects of Admission on Program Participation, 2012–2020



(a) Afternoon



(b) Pull-out



(c) Full-time

Notes: Each panel plots observed program participation against admission-test score points relative to the admission cutoff for cohorts tested between 2012 and 2020. The discontinuity measures how crossing the admission cutoff affects participation in each program. These later cohorts are used to assess take-up because direct participation measures are incomplete for the older cohorts used in the long-run analysis.

3.1 Participation

Figure 2 shows the effect of admission on participation between 2012–2020. Overall, we see a large effect on participation in all three programs. However, there are important distinctions between the programs. We discuss these distinctions starting with the simplest case: the afternoon program.

Panel (a) of Figure 2 shows the increase in participation following admission to the afternoon program. Approximately 20% of students just below the admission threshold still find their way into the program. We estimate that passing the admission threshold increases this share by 46 percentage points (p.p.). On both sides of the threshold, the share of students who participate in one of the other programs is virtually zero. This is expected, as the admission threshold to the afternoon program is lower.

Panel (b) shows that admission to the pull-out program increases participation in that

program by 85 p.p. With a few rare exceptions, students who were not admitted do not participate in the pull-out program. However, students who narrowly missed admission can still enroll in the afternoon program where it is available. About 50% of these students participate. Therefore, for many students the effect of admission to the pull-out program represents the effect of switching from the afternoon to the pull-out program.

Panel (c) shows that admission to the full-time program increases participation in this program by 46 p.p. This effect is similar to the effect of the afternoon program and lower than the effect of admission to the pull-out program. We believe this difference reflects the larger commitment required by the full-time program, as students leave their original class and neighborhood school altogether and may face a substantial daily commute. As in the pull-out program, students below the cutoff cannot participate in the full-time program but are still entitled to participate in the afternoon program, where it is available. Yet, less than 10% do so. Finally, some students admitted to the full-time program eventually participate in the pull-out program. This could be in the same municipality, or in pull-out programs in a nearby municipality.[11] Therefore, the effect of admission to the full-time program combines effects of participation in both the full-time and pull-out programs.

4 Private Benefits: Educational and Labor-Income Effects

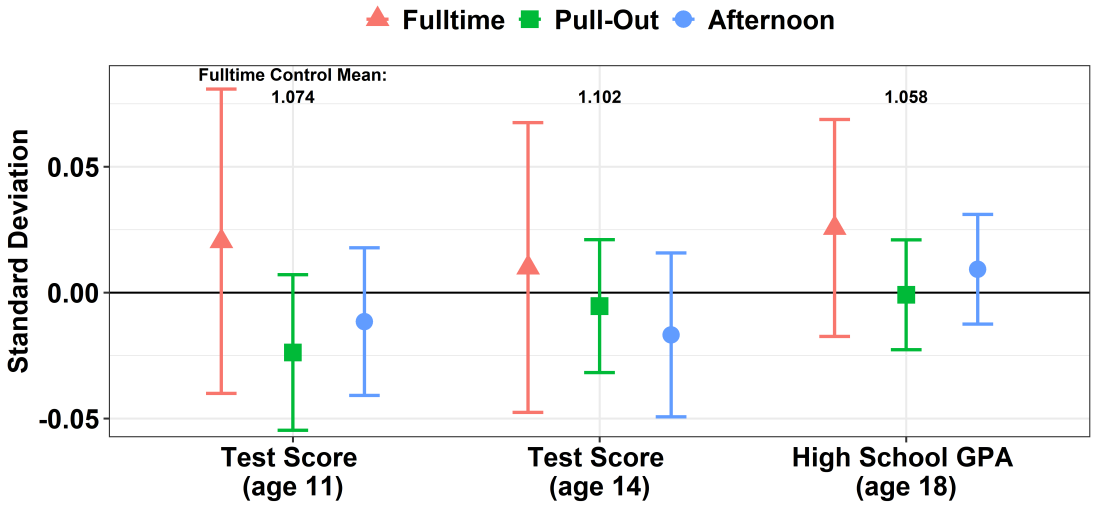
4.1 Short-Term Effects

We first examine the effects on standardized national test scores in grades 5 and 8 and on high-school GPA. These outcomes are closest to those emphasized in the existing literature on gifted programs (e.g., Bui et al., 2014; Card and Giuliano, 2014; Chin and Kocks, 2025). Consistent with the literature, we find effects close to zero for all programs at all ages.

Figure 3 plots the point estimates from the RDD of the effect of admission on test scores at different ages. Each symbol and color represent the three program types: full-time (red triangles), pull-out (green squares), and afternoon (blue dots). Test scores are measured in standard deviation units.

[11] For older cohorts, the pull-out program was only available through a neighboring municipality.

Figure 3: Short- and Medium-Run Educational Outcomes



Notes: The figure reports admission effects on standardized test scores around age 11 (grade 5), standardized test scores around age 14 (grade 8), and standardized high-school GPA around age 18. The age-11 test-score outcome averages standardized scores in mathematics, Hebrew or Arabic, and English; the age-14 outcome additionally includes science and technology. Estimates are shown separately for the full-time, pull-out, and afternoon programs. Markers are point estimates and vertical bars are 90 percent confidence intervals.

While test scores are often useful proxies for later-life achievements, the programs we study were not designed to improve them. Moreover, since students in the control group (just below the admission cutoff) already have very high test scores, there is little room for further improvement. For these reasons, the absence of test-score effects does not necessarily extend to longer-term outcomes. And indeed, as we discuss next, the picture changes considerably when we turn to these outcomes.

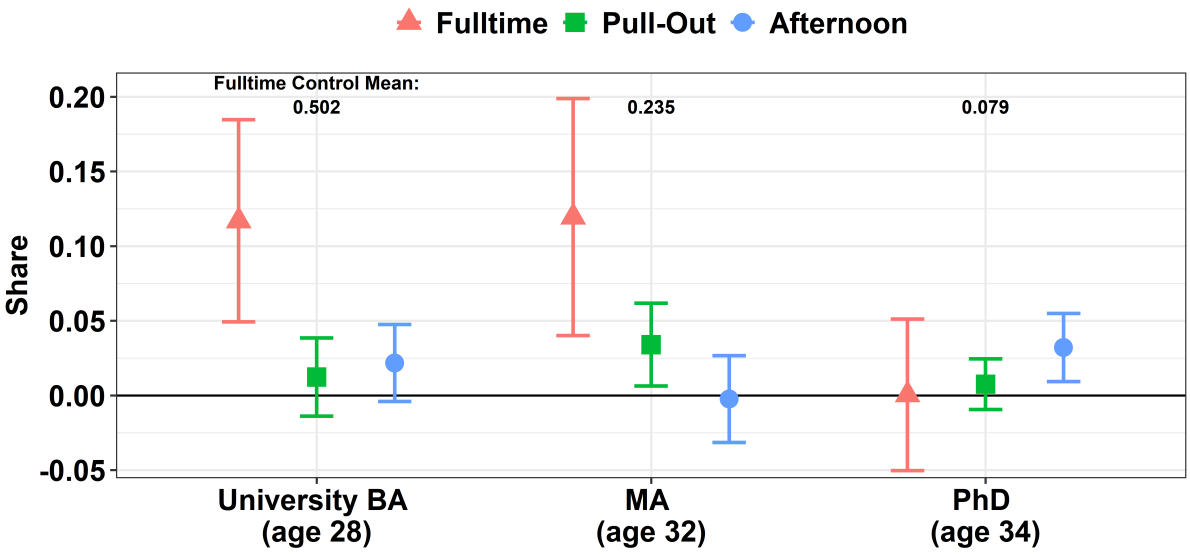
4.2 Long-Term Education Effects

In this section we consider effects of the programs on long-term educational outcomes. We begin by analyzing the effect on the attainment of various academic degrees.

Figure 4 plots the estimates from the RDD of the effect of admission on university bachelor's-degree (BA) attainment by age 28, master's-degree (MA) attainment by age 32, and PhD attainment by age 34.[12] We find substantial effects, with the strongest effects in the full-time program. Admission to the full-time program increases the probability of earning a university BA by age 28 by 12 p.p., relative to a baseline of 50 percent. It also increases

[12]We chose ages based on the median age of graduates in each degree level.

Figure 4: Effects on Higher-Education Attainment

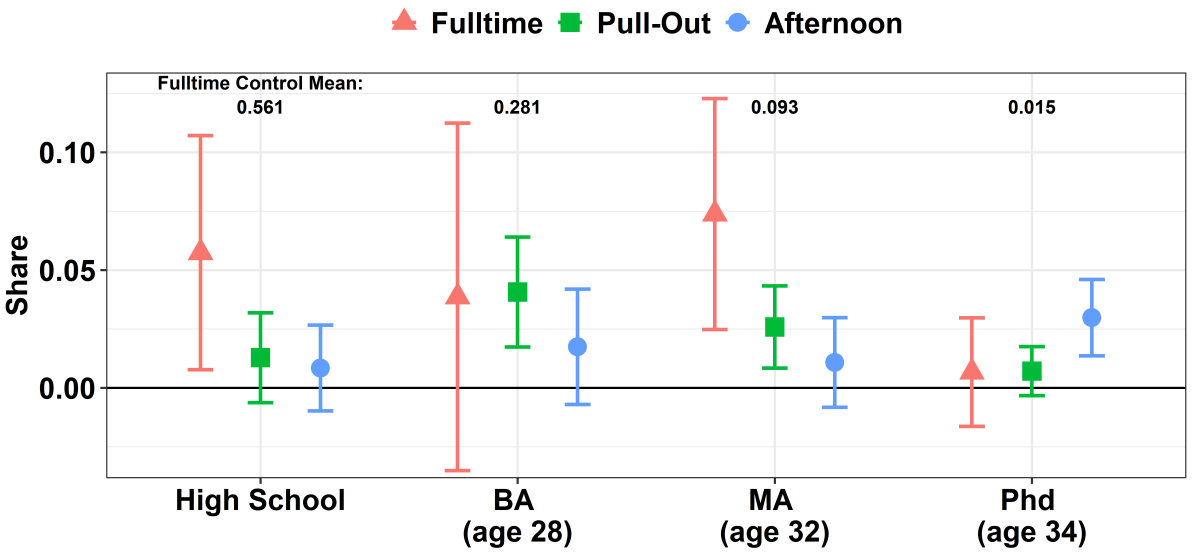


Notes: The figure reports admission effects on bachelor's-degree attainment by age 28, master's-degree attainment by age 32, and PhD attainment by age 34. Estimates are shown separately for the three programs. Each regression uses applicants old enough to observe the listed outcome. Markers are point estimates and vertical bars are 90 percent confidence intervals.

MA attainment by 12 p.p., relative to a baseline of 24 percent, a 50% increase. While we do not find a statistically significant effect of the full-time program on PhD attainment, the estimates are imprecise enough that we cannot rule out a proportional increase of a similar magnitude. We also find a 3 p.p. increase in MA attainment following admission to the pull-out program and a 3 p.p. increase in PhD attainment following admission to the afternoon program.

Figure 5 shows that the effects are particularly concentrated in STEM fields. For the full-time program, these effects are already apparent in high school: admission to the full-time program increase the probability of advanced STEM study by 6 p.p., from a baseline of 56 percent. It also increases the probability of completing a STEM MA by 7 p.p., from a baseline of 9 percent. We find positive but not statistically significant effects on STEM BA and PhD attainment. The pull-out program increases the probability of acquiring a STEM BA and MA by 4 and 3 p.p., respectively. The afternoon program also increases PhD attainment by 3 p.p. We do not find similar effects on non-STEM degrees.

Figure 5: Effects on STEM Education



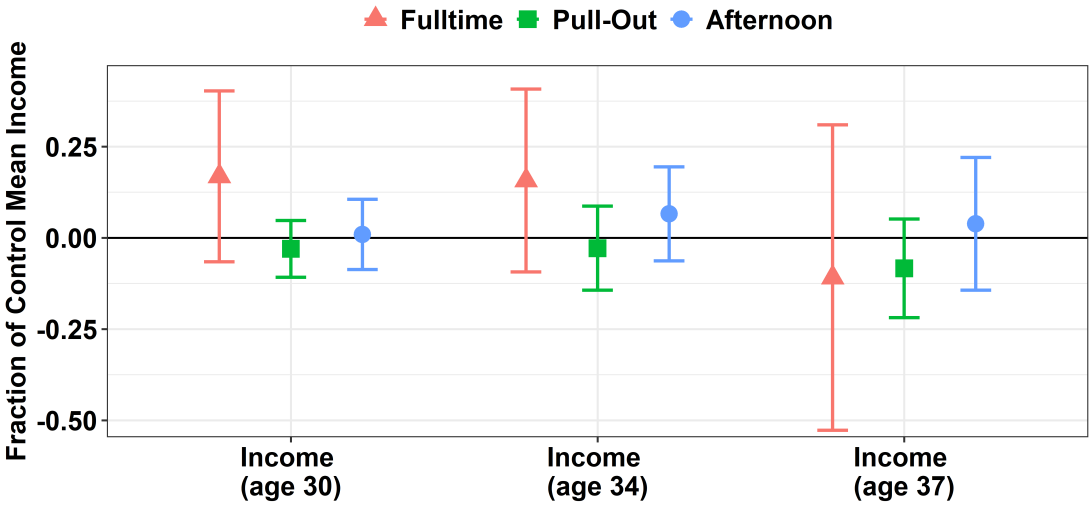
Notes: The figure reports the effects of admission on advanced STEM study in high school and STEM-degree attainment at the undergraduate and graduate levels, separately for the full-time, pull-out, and afternoon programs. Markers are point estimates and vertical bars are 90 percent confidence intervals. Each regression uses applicants old enough to observe the listed outcome.

4.3 Labor Income

Figure 6 reports effects on labor income at ages 30, 34, and 37, scaled by the corresponding control mean. The point estimates generally follow the education results, at least for the full-time program, but they are less precise. While we cannot rule out null effects, we also cannot rule out very substantial effects. There are at least two reasons for this imprecision. First, labor income is noisier than degree attainment. Second, even age 37 may still be early for this highly educated population in Israel because compulsory military service and lengthy undergraduate and graduate study delay entry into stable careers.

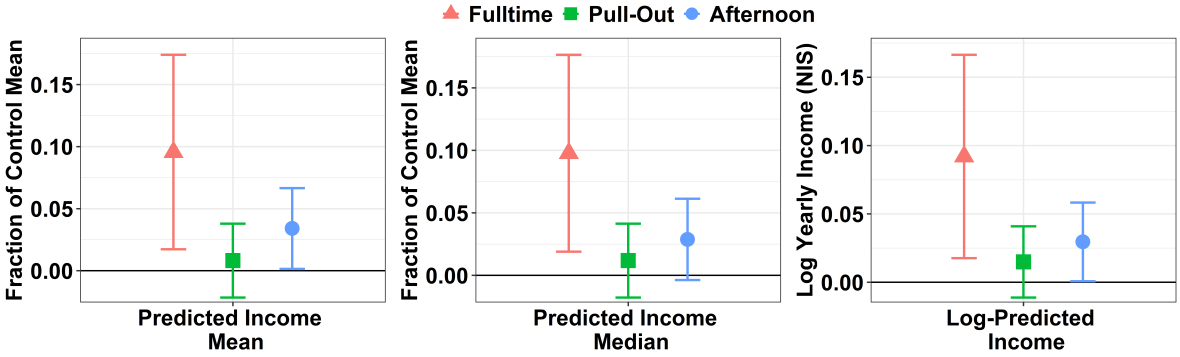
To assess whether the educational changes are likely to translate into later labor income, we also construct predicted age-37 labor income using earlier degree characteristics, including degree field and institution. Figure 7 reports the results that display estimates that are more similar to the education patterns.The estimates predict an effect of about 10 p.p increase in income by age 37 for those admitted to the full-time program. The effect of admission to the pull-out program is not statistically significant. Admission to the afternoon program is predicted to raise income by about 3 p.p.

Figure 6: Effects on Labor Income by Age



Notes: The figure reports the effects of admission on total reported labor income at ages 30, 34, and 37. Markers are point estimates and vertical bars are 90 percent confidence intervals. Each regression uses applicants old enough to observe the listed outcome.

Figure 7: Effects on Predicted Later-Life Labor Income

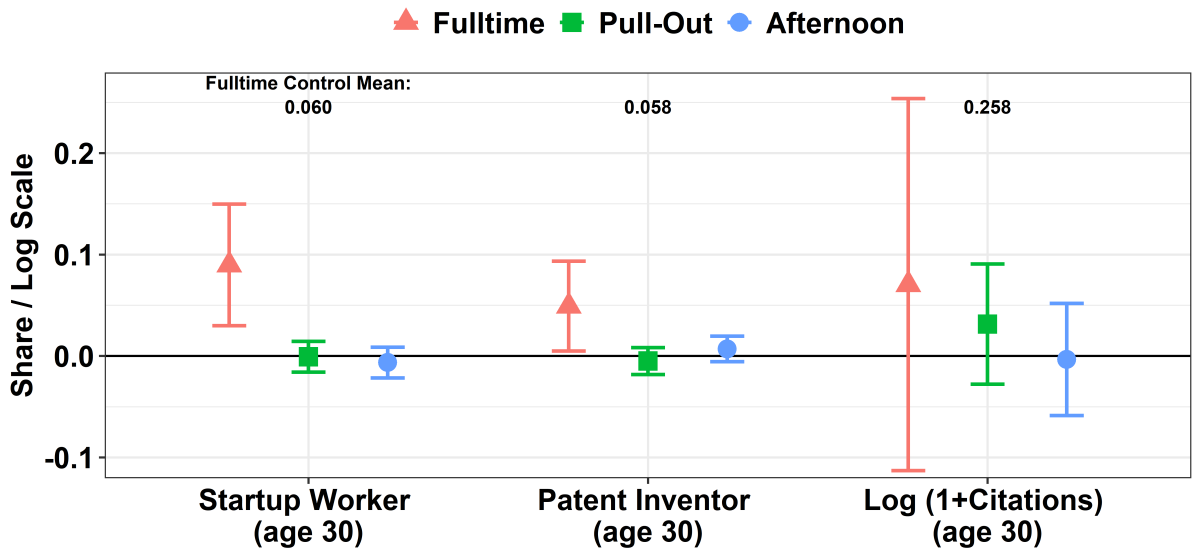


Notes: The figure reports the effects of admission on predicted later-life labor income, based on earlier educational outcomes, including degree field and institution. The predicted income reflects the labor-market returns associated with these educational choices.

5 Innovation

A central justification for gifted education is that benefits may extend beyond participants through innovation and the production of knowledge. To test this, we follow previous literature and examine three sets of outcomes. First, we measure inventive activity using patenting (Trajtenberg, 1990; Bell et al., 2019). Second, we examine employment in startups and the technology sector, which play an important role in developing and commercializing new technologies (Kortum and Lerner, 2000; Akcigit and Kerr, 2018). Third, we measure scientific

Figure 8: Effects on Startups and Patenting



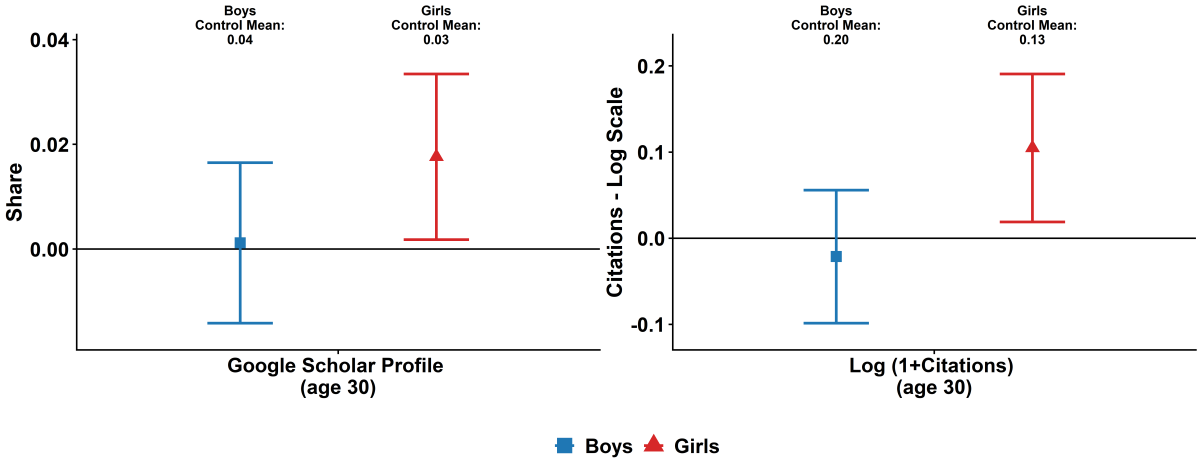
Notes: This figure reports the effects of program admission on startup employment by age 30, patent inventor status by age 30, and log(1 + patent citations) by age 30. Patent outcomes come from United States Patent and Trademark Office records accessed through PatentsView. Markers are point estimates and vertical bars are 90 percent confidence intervals. Each regression uses applicants old enough to observe the listed outcome.

activity using whether an individual has a Google Scholar profile and the number of citations received (Jaffe, 1989; Adams, 1990).

Figure 8 shows strong positive effects on two out of the three innovation outcomes. Admission increases the chances of being a patent inventor by 5 p.p., almost doubling the baseline rate. Admission increases the probability of working in a startup and the probability of being identified as an inventor. It also increases patent citations for some programs, though these citation effects are not statistically significant. The patterns broadly mirror the STEM-education results, suggesting that the programs affect not only credentials but also later activities closely related to innovation.

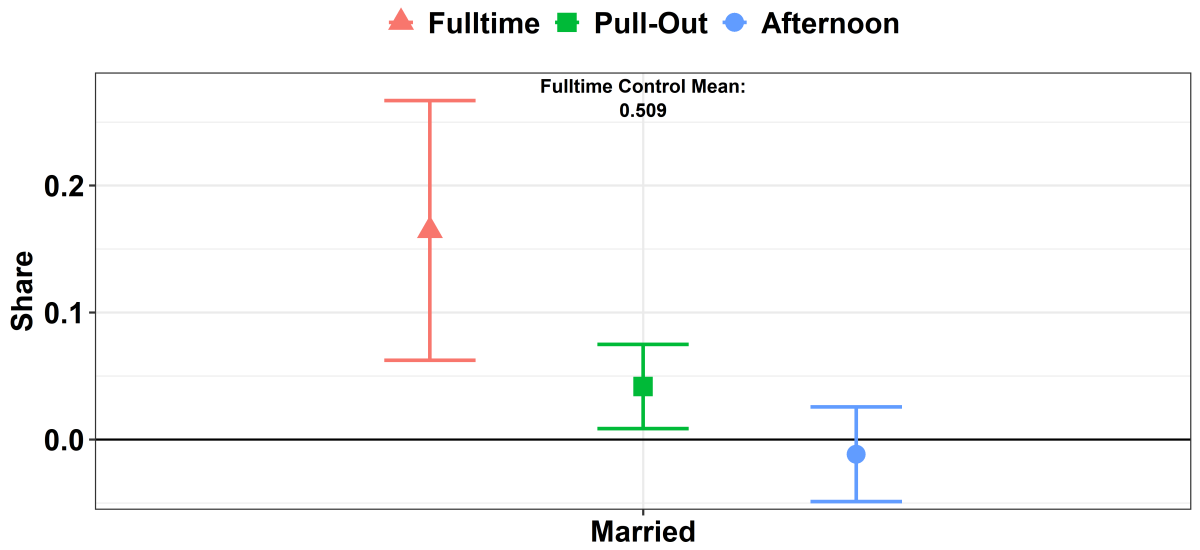
The evidence on academic publications is less precise. Figure 9 shows positive Google Scholar effects for women in the pull-out program, while we do not find corresponding effects for the full-time or afternoon programs. We interpret these results together with the patent and startup evidence.

Figure 9: Effects on Academic Publications in the Pull-Out Program, by Gender



Notes: The figure reports pooled RDD estimates of admission to the pull-out program on having an academic-publication profile and on $\log(1 + \text{citations})$, separately for men and women. Markers are point estimates and vertical bars are 90 percent confidence intervals. Each regression uses applicants old enough to observe the listed outcome.

Figure 10: Effects on Marriage



Notes: The figure reports the effects of program admission on ever being married by age 30 among applicants old enough to be observed through age 30. Estimates are shown separately for the full-time, pull-out, and afternoon programs. Marriage is measured using the population registry. Markers are point estimates and vertical bars are 90 percent confidence intervals.

6 Marriage

Figure 10 shows that program admission increases the probability of marriage by age 30. Admission to the full-time program increases the marriage rate by 16 p.p., relative to a baseline of 51%. Admission to the pull-out program also generates a smaller, but statistically

significant increase of 4 p.p.

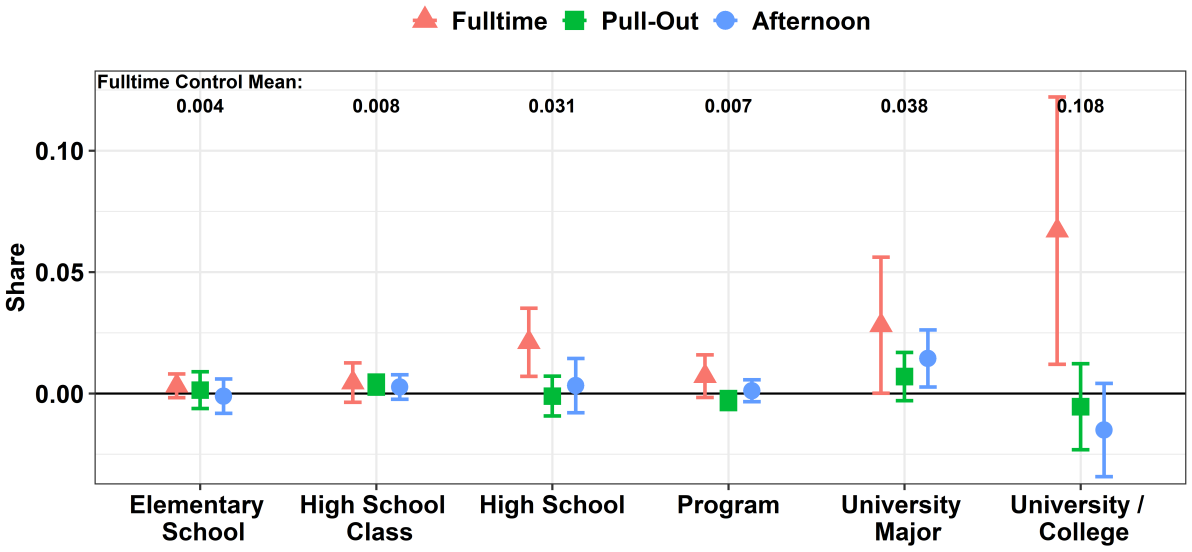
The marriage effects partly close a pre-existing gap between gifted students and the general population. The summary statistics in Table 2 show that admitted students are less likely to be married by age 30 than the general population. Program admission narrows this difference. This pattern contrasts with our previous results, where for most other outcomes, the programs amplify pre-existing differences between gifted students and the broader population by further improving outcomes on which gifted students already tend to perform well.

We consider two possible mechanisms for the increase in marriage. First, admission may change students' educational trajectories and thereby the pool of potential partners they encounter. As we showed in Section 4.2, admission to the programs substantially affects later-life educational paths. Figure 11 examines whether admitted students are more likely to marry someone with whom they share one of the educational environments listed in the figure. Admission to the full-time program increases the probability of marrying someone who attended the same high school, the same major in the university, or the same university more broadly. This is consistent with the previous literature in which effects of educational policies on marriage are more commonly documented for interventions operating in high school or higher education. (Currie and Moretti, 2003; Krdar et al., 2018; Yang, 2022; Kirkebøen et al., 2025).

Second, the programs may develop interpersonal or socio-emotional skills that increase the probability of forming a marriage. For instance, one of the few well-known examples of an early-childhood intervention with long-run effects on marriage is Perry Preschool, which also produced substantial effects on socio-emotional skills (Heckman et al., 2013; García et al., 2023).

To shed further light on these mechanisms, we examine the academic ability of the spouses whom admitted students marry. We measure the academic ability of the spouse using the University Psychometric Entrance Test (UPET) score. We divide spouses into three groups based on their UPET scores: below the population average, between the population average and one standard deviation above, and more than one standard deviation above the population average. We also consider a fourth ability group consistent of spouses who were themselves admitted to one of the three programs we study.

Figure 11: Effects on Marrying Someone From a Shared Educational Environment

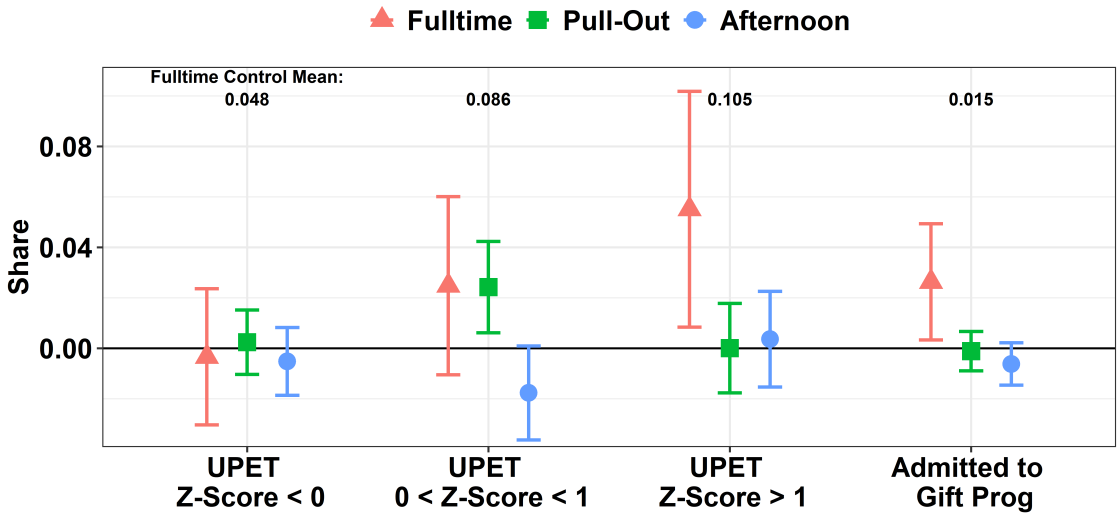


Notes: This figure reports the effects of admission on ever marrying someone who shared each educational environment listed on the horizontal axis, for individuals above age 25. Markers are point estimates and vertical bars are 90 percent confidence intervals. Elementary school includes individuals who overlapped in the same elementary school. High school class includes students in the same homeroom class in high school, while high school includes any overlap in high school. Program includes students who participated in the same gifted program. University major includes all students who studied the same university major at the same institution during the same period, while university/college includes any overlap in the same institution.

Figure 12 examines the effect of admission on marriage with a spouse from each of these groups. Even among students just below the admission cutoff, marriages are strongly positively selected on spouse ability. For students just below the cutoff of the full-time program, fewer than 5% marry spouses with below-average UPET scores, compared with more than 10% who marry spouses scoring at least one standard deviation above-average UPET. Admission to the full-time program further increases the probability of marrying high-ability spouses: the effect of marrying a spouse with one standard deviation above-average score is 6 p.p., and the effect of marrying a student who was admitted to one of the three programs is 3 p.p., from a baseline of 1.5%. This is consistent with the results in Figure 11 that show an effect of the full-time program on marrying from the same education program, which is likely very selective as well.

By contrast, the pull-out program only increases marriage for the middle group of spouses whose UPET lies between the population average and one standard deviation above it. Thus, admission to the pull-out program is not shifting marriages toward especially high-ability

Figure 12: Effects on Spouse University Entrance-Test Scores



Notes: The figure reports the effect of admission on marriage with spouses from a particular ability group. We classify spouses into three groups based on their standardized (Z-score) university psychometric entrance test (UPET). The fourth group includes all spouses who were admitted to one of the three programs we study (full-time, pull-out, afternoon). Markers are point estimates and vertical bars are 90 percent confidence intervals.

spouses. Instead, it appears to increase the probability of marrying spouses whose observed academic ability is similar to that of the typical spouse in the control group. This is less consistent with greater exposure to high-ability peers, and more consistent with a general increase in the probability of marriage. One possible explanation is that the program improves interpersonal or socio-emotional skills that are valuable in forming a marriage relationship.

7 Impact on Other Students

Critics of gifted student programs worry that removing gifted students from their neighborhood homerooms may negatively affect the students who remain. The concern is most relevan for the full-time program, which moves participant students out of their original homeroom. It also applies, though to a lesser extent, to the pull-out program, which removes participants from their homeroom class one day a week. The afternoon program does not remove students during school hours, but it may still affect classmates through changes in the admitted students' knowledge, behavior, or aspirations. It could also reduce the time and social interaction that participants have with their original homeroom classmates if friendships formed in the afternoon program substitute for those relationships.

Table 5: Effects on Second-Grade Classmates

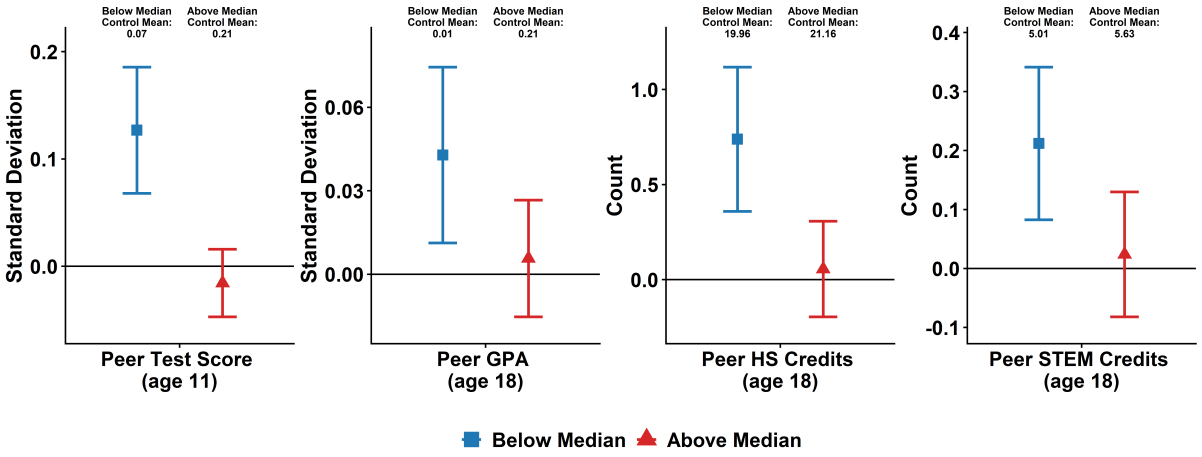
	Full-time			Pull-out			Afternoon		
Outcome	Mean	(1)	(2)	Mean	(3)	(4)	Mean	(5)	(6)
Peer grade-5 test score (std.)		0.01	0.01		0.00	-0.00		0.01	0.01
		(0.03)	(0.03)		(0.02)	(0.02)		(0.02)	(0.02)
	0.31	[3,894]	[3,673]	0.15	[19,685]	[19,882]	0.17	[22,704]	[21,660]
Peer high-school GPA (std.)		-0.02	-0.01		0.00	-0.01		0.01	0.01
		(0.03)	(0.02)		(0.01)	(0.01)		(0.02)	(0.01)
	0.31	[6,685]	[6,591]	0.17	[34,160]	[35,437]	0.14	[40,674]	[34,822]
Peer high-school credits		-0.13	-0.05		0.11	0.02		0.13	0.24
		(0.29)	(0.26)		(0.15)	(0.12)		(0.16)	(0.15)
	21.58	[5,985]	[5,202]	21.10	[35,064]	[39,257]	20.84	[36,762]	[33,124]
Peer advanced mathematics		-0.00	0.00		0.00	0.00		0.00	0.00
		(0.01)	(0.01)		(0.00)	(0.00)		(0.00)	(0.00)
	0.22	[6,906]	[5,811]	0.18	[31,494]	[30,080]	0.18	[40,084]	[40,264]
Peer MA		-0.00	-0.00		-0.00	-0.00		0.00	0.00
		(0.02)	(0.02)		(0.01)	(0.01)		(0.01)	(0.01)
	0.11	[882]	[836]	0.09	[7,625]	[7,962]	0.09	[6,904]	[7,090]
Peer started PhD by 32		-0.00	-0.00		-0.00	-0.00		0.00	0.00
		(0.01)	(0.01)		(0.00)	(0.00)		(0.00)	(0.00)
	0.03	[881]	[945]	0.02	[6,942]	[6,855]	0.02	[5,414]	[5,467]
Controls		No	Yes		No	Yes		No	Yes

Notes: Classmates are students assigned to the applicant's second-grade homeroom, excluding the applicant. Each coefficient is a pooled RDD estimate of how the applicant's crossing of the admission cutoff affects the listed outcome among classmates. Standard errors are in parentheses and the number of observations is in square brackets. Columns labeled "Yes" include the study's predetermined controls.

We identify classmates using the applicant's second-grade homeroom, before the admission decision. We estimate how the applicant's crossing of the admission cutoff affects the average outcomes of the other students in that homeroom, excluding the applicant. The estimates therefore capture the effect generated by admitting one marginal student, averaged across the classmates linked to each applicant.

We find that the effects of an admitted classmate are very close to zero. Table 5 reports effects on classmates' grade-5 test scores, high-school GPA and credits, advanced mathematics, and graduate-degree outcomes. We report estimates with and without control variables. In all specifications and all programs, the estimates are very close to zero and with tight confidence intervals. For example, the effect on fifth grade test scores is between -.01 and .01 standard deviations, with standard errors around .02-.03. This table rules out the large adverse effects sometimes invoked in the policy debate.

Figure 13: Educational Effects on Classmates of Afternoon-Program Applicants, by Socioeconomic Status



Notes: The figure reports RDD estimates of admission to the afternoon program on classmates' grade-5 test scores, high-school GPA, total high-school credits, and STEM credits, separately by socioeconomic status. Classmates are identified from the applicant's second-grade homeroom. Because the afternoon program occurs outside regular school hours, the estimates capture spillovers from the admitted student's participation rather than the effect of removing that student from class.

7.1 Heterogeneity by Socioeconomic Status

Average effects may conceal different spillovers across school environments. Figure 13 shows heterogeneity in afternoon-program spillovers by the socioeconomic status of the classmates. Following the ordering in the figure, admission has positive effects on below-median classmates' grade-5 test scores, high-school GPA, total high-school credits, and STEM credits, while the corresponding estimates for above-median classmates are close to zero. A possible mechanism is that admitted students transmit knowledge, information, or aspirations to their homeroom classmates—for example, while studying or doing homework together.

The result should be interpreted cautiously. The afternoon program does not remove the admitted student during school hours, so it is not a direct test of the cost of removing a high-ability peer. Moreover, the mechanism is not identified. Nevertheless, the positive effects on low-SES classmates reinforce the conclusion that the afternoon program does not generally impose negative spillovers on classmates from different socioeconomic backgrounds.

8 Conclusion

Whether governments should invest in programs for gifted students depends on outcomes that extend beyond immediate test scores. The programs impose direct fiscal costs on governments and private costs on participating families. They may also create indirect costs if they increase inequality or harm other students. Against these costs, the programs may generate gains to participants, support students with distinctive developmental needs, and produce external benefits through innovation. Existing evidence has not resolved this debate because it has typically focused on short-run academic outcomes.

We provide causal evidence on the long-run effects of three national elementary-school programs in Israel. The programs differ in intensity, timing, and private cost, but share a centralized admission examination and enrichment curriculum. Exploiting local admission cutoffs, we find that short-run effects on test scores are zero or negative. These results are similar to the existing literature. In the long run, however, admission increases higher-education attainment, STEM education, and several measures of innovation. Admission also increases marriage rates. The full-time program generally has the largest effects, suggesting that intensity and sustained exposure to high-ability peers may matter.

The distributional results are mixed but informative. In the older cohorts studied here, girls and students from lower socioeconomic backgrounds were underrepresented among admitted students. Yet women often experience larger treatment effects. The historical admission process therefore left potential gains unrealized among groups that were less likely to participate. Policies that improve screening, reduce tuition and transportation barriers, and broaden access could increase both equity and aggregate benefits.

We also find no evidence of substantial negative effects on classmates. Average spillovers are close to zero, and the afternoon program generates positive educational spillovers in lower-SES settings. These findings do not imply that every expansion is costless: a large-scale expansion could affect teacher supply, program quality, and classroom composition in ways that differ from the marginal admissions studied here. They do show, however, that the existing programs generate benefits that are not visible in short-run test scores and do not appear to do so at the expense of classmates.

A complete policy evaluation requires program costs and evidence on external validity beyond the current cutoffs. The comparison across the three programs suggests that full-time

programs produce larger effects, while afternoon programs can generate meaningful benefits at lower intensity. Future work should combine the causal estimates with detailed cost data to compare universal expansion with targeted expansion in places where access to advanced STEM instruction is limited.

References

Abdulkadiroglu, A., J. Angrist, and P. Pathak (2014). "The Elite Illusion: Achievement Effects at Boston and New York Exam Schools". In: *Econometrica* 82, pp. 137–196.

Adams, J. D. (Aug. 1990). "Fundamental Stocks of Knowledge and Productivity Growth". In: *Journal of Political Economy* 98.4, pp. 673–702.

Akcigit, U. and W. R. Kerr (2018). "Growth through Heterogeneous Innovations". In: *Journal of Political Economy* 126.4, pp. 1374–1443.

Athey, S., R. Chetty, G. W. Imbens, and H. Kang (2026). "The surrogate index: Combining short-term proxies to estimate long-term treatment effects more rapidly and precisely". In: *Review of Economic Studies* 93.4, pp. 2284–2312.

Balestra, S., A. Sallin, and S. C. Wolter (2023). "High-Ability Influencers? The Heterogeneous Effects of Gifted Classmates". In: *Journal of Human Resources* 58.2, pp. 633–665.

Bell, A., R. Chetty, X. Jaravel, N. Petkova, and J. Van Reenen (2019). "Who becomes an inventor in America? The importance of exposure to innovation". In: *The Quarterly Journal of Economics* 134.2, pp. 647–713.

Bernstein, B. O., D. Lubinski, and C. P. Benbow (2021). "Academic Acceleration in Gifted Youth and Fruitless Concerns Regarding Psychological Well-Being: A 35-Year Longitudinal Study". In: *Journal of Educational Psychology* 113.4, pp. 830–845.

Boisjoly, J., G. J. Duncan, M. Kremer, D. M. Levy, and J. Eccles (2006). "Empathy or antipathy? The impact of diversity". In: *American Economic Review* 96.5, pp. 1890–1905.

Booij, A. S., F. Haan, and E. Plug (2016). *Enriching students pays off: Evidence from an individualized gifted and talented program in secondary education*. Tech. rep. IZA Discussion Papers.

Borland, J. H. (2003). "The death of giftedness: Gifted education without gifted children". In: *Rethinking gifted education* 10, pp. 105–124.

Bui, S. A., S. G. Craig, and S. A. Imberman (2014). "Is gifted education a bright idea? Assessing the impact of gifted and talented programs on students". In: *American Economic Journal: Economic Policy* 6.3, pp. 30–62.

Burris, C. C. and D. T. Garrity (2008). *Detracking for excellence and equity*. ASCD.

Card, D., E. Chyn, and L. Giuliano (2026). "Can Gifted Education Help Higher-Ability Boys from Disadvantaged Backgrounds?" In: *Journal of Political Economy*. Forthcoming.

Card, D. and L. Giuliano (2014). "Does gifted education work? For which students?" In: *National Bureau of Economic Research*.

— (2016a). "Can tracking raise the test scores of high-ability minority students?" In: *American Economic Review* 106.10, pp. 2783–2816.

— (2016b). "Universal screening increases the representation of low-income and minority students in gifted education". In: *Proceedings of the National Academy of Sciences* 113.48, pp. 13678–13683.

Cattaneo, M. D., R. Titiunik, G. Vazquez-Bare, and L. Keele (2016). "Interpreting regression discontinuity designs with multiple cutoffs". In: *The Journal of Politics* 78.4, pp. 1229–1248.

Chin, J. and G. Kocks (Nov. 2025). "Sorting or Supporting? The Effect of Gifted Education on Achievement and Access". Job Market Paper.

Cleveland, C. H. (2023). "Gifted or Gone: The Impacts of Gifted Education on Public School Enrollment and Achievement". Working paper.

Cohodes, S. R. (2020). "The long-run impacts of specialized programming for high-achieving students". In: *American Economic Journal: Economic Policy* 12.1, pp. 127–166.

Currie, J. and E. Moretti (2003). "Mother's Education and the Intergenerational Transmission of Human Capital: Evidence from College Openings". In: *The Quarterly Journal of Economics* 118.4, pp. 1495–1532.

Danieli, O., D. Nevo, I. Walk, B. Weinstein, and D. Zeltzer (2024). "Negative Control Falsification Tests for Instrumental Variable Designs". In: *working paper*.

Davis, B., J. Engberg, D. N. Epple, H. Sieg, and R. Zimmer (2010). *Evaluating the gifted program of an urban school district using a modified regression discontinuity design*. Tech. rep. National Bureau of Economic Research.

Dobbie, W. and R. G. Fryer (2014). "The impact of attending a school with high-achieving peers: Evidence from the New York City exam schools". In: *American Economic Journal: Applied Economics* 6, pp. 58–75.

Freeman, J. (2006). "Giftedness in the Long Term". In: *Journal for the Education of the Gifted* 29.4, pp. 384–403.

García, J. L., J. J. Heckman, and V. Ronda (2023). "The Lasting Effects of Early-Childhood Education on Promoting the Skills and Social Mobility of Disadvantaged African Americans and Their Children". In: *Journal of Political Economy* 131.6, pp. 1477–1506.

Goldstein, Y., V. Lavy, and G. Rachkovski (Apr. 2026). *Gifted Children Programs' Short- and Long-Term Impact: Higher Education, Earnings, and the Knowledge Economy*. NBER Working Paper 29779. National Bureau of Economic Research.

Gross, M. U. M. (2006). "Exceptionally Gifted Children: Long-Term Outcomes of Academic Acceleration and Nonacceleration". In: *Journal for the Education of the Gifted* 29.4, pp. 404–429.

Heckman, J. J., J. Stixrud, and S. Urzua (2006). "The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior". In: *Journal of Labor economics* 24.3, pp. 411–482.

Heckman, J. J., R. Pinto, and P. A. Savelyev (2013). "Understanding the Mechanisms through Which an Influential Early Childhood Program Boosted Adult Outcomes". In: *American Economic Review* 103.6, pp. 2052–2086.

Helal, A. M., F. A. Parra-Marinez, R. R. Djita, J. Wai, and S. McKenzie (2026). "Gifted Students' Later Outcomes: College-Going Rates and Selectivity". In: *Gifted Child Quarterly* 70.1, pp. 3–19.

Hoisl, K., H. C. Kongsted, and M. Mariani (2023). "Lost Marie Curies: Parental impact on the probability of becoming an inventor". In: *Management Science* 69.3, pp. 1714–1738.

Jaffe, A. B. (Dec. 1989). "Real Effects of Academic Research". In: *American Economic Review* 79.5, pp. 957–970.

Kirkebøen, L., E. Leuven, M. Mogstad, and J. Mountjoy (2025). *College as a Marriage Market*. Tech. rep. 25-1199. Annenberg Institute at Brown University.

Krdar, M. G., M. Dayolu, and. Koç (2018). "The Effects of Compulsory-Schooling Laws on Teenage Marriage and Births in Turkey". In: *Journal of Human Capital* 12.4, pp. 640–668.

Kline, P. and C. R. Walters (2016). "Evaluating public programs with close substitutes: The case of Head Start". In: *The Quarterly Journal of Economics* 131.4, pp. 1795–1848.

Kortum, S. S. and J. Lerner (2000). "Assessing the Contribution of Venture Capital to Innovation". In: *The RAND Journal of Economics* 31.4, pp. 674–692.

Lundberg, S. (2012). "Personality and Marital Surplus". In: *IZA Journal of Labor Economics* 1.1, p. 3.

Mammadov, S. and N. B. Hertzog (2024). "Million Dollar Question: What Is the Most Effective and Equitable Way to Deliver Services to Advanced Learners?" In: *Journal of School Leadership* 34.2, pp. 151–176.

McCrary, J. (2008). "Manipulation of the running variable in the regression discontinuity design: A density test". In: *Journal of econometrics* 142.2, pp. 698–714.

Moreira, D. (2016). "Success spills over: How awards affect winners and peers performance in Brazil". In: *Manuscript*.

Noy, M. and Y. Vorgan (July 2026). *Identification and Development of Gifted Students in the Education System [in Hebrew]*. Review. Hebrew-language report; title translated into English. Jerusalem, Israel: Knesset Research and Information Center.

Olszewski-Kubilius, P., S.-Y. Lee, M. Ngoi, and D. Ngoi (2004). "Addressing the achievement gap between minority and nonminority children by increasing access to gifted programs". In: *Journal for the Education of the Gifted* 28.2, pp. 127–158.

Plucker, J. A., N. Burroughs, and R. Song (2010). "Mind the (Other) Gap! The Growing Excellence Gap in K-12 Education." In: *Center for Evaluation and Education Policy, Indiana University*.

Redding, C. and J. A. Grissom (2021). "Do students in gifted programs perform better? Linking gifted program participation to achievement and nonachievement outcomes". In: *Educational Evaluation and Policy Analysis* 43.3, pp. 520–544.

Renzulli, J. (2021). "The major goals of gifted education and talent development programs". In: *Academia Letters* 2585.

Rogers, K. B. (2007). "Lessons Learned About Educating the Gifted and Talented: A Synthesis of the Research on Educational Practice". In: *Gifted Child Quarterly* 51.4, pp. 382–396.

Rutigliano, A. and N. Quarshie (2021). "Policy approaches and initiatives for the inclusion of gifted students in OECD countries". In: *OECD Education Working Papers* 262, 0_1–71.

Silverman, L. K. (2017). "The construct of asynchronous development". In: *Charting a new course in gifted education*. Routledge, pp. 36–58.

Steenbergen-Hu, S., M. C. Makel, and P. Olszewski-Kubilius (2016). "What One Hundred Years of Research Says About the Effects of Ability Grouping and Acceleration on K–12 Students' Academic Achievement: Findings of Two Second-Order Meta-Analyses". In: *Review of Educational Research* 86.4, pp. 849–899.

Subotnik, R. F., P. Olszewski-Kubilius, and F. C. Worrell (2011). "Rethinking Giftedness and Gifted Education: A Proposed Direction Forward Based on Psychological Science". In: *Psychological Science in the Public Interest* 12.1, pp. 3–54.

Thompson, O. (2024). "Gifted & Talented Programs and Racial Segregation". In: *Education Finance and Policy* 19.4, pp. 692–715.

Trajtenberg, M. (1990). "A Penny for Your Quotes: Patent Citations and the Value of Innovations". In: *The RAND Journal of Economics* 21.1, pp. 172–187.

Wai, J., D. Lubinski, C. P. Benbow, and J. H. Steiger (2010). "Accomplishment in Science, Technology, Engineering, and Mathematics (STEM) and Its Relation to STEM Educational Dose: A 25-Year Longitudinal Study". In: *Journal of Educational Psychology* 102.4, pp. 860–871.

Yang, S. (2022). "The Effects of Compulsory Schooling Reforms on Women's Marriage Outcomes—Evidence from Britain". In: *Empirical Economics* 63.3, pp. 1637–1662.

Appendix

A Appendix Figures and Tables

Figure A.1: Example Schedule in a Pull-Out Program

Classsrom	C. 1	C. 2	C. 3	C. 4	C. 5	C. 6	C. 7	C. 8	C. 9	C. 10	C. 11
Unit 1	Numbers and games	Theater	Food for the body and soul	The wonderful world of mathematics	Math from a different angle	The inventions that changed the world	Art workshop following wall paintings	The chemistry of everyday	Dungeons and Dragons	The wonderful numbers	Animation
8:15-9:45	Veronica	Neta	Efrat	Gilad	Jeff	Yvonne	Ruthie	Michal	Ran	Liat	Nurit
Snack 9:45-9:45											
Break 9:55-10:20											
Unit 2	Suns, comets and stars	Numbers and games	``Fantastic Beasts and Where to Find Them'' ecology	Food for the body and soul	The soul bird	Animation	Dungeons and Dragons	Art workshop following wall paintings	Math from a different angle	Mechanics and chain reactions	Tactical-strategic games
10:20-11:40	Meir P.	Veronica	Dana	Efrat	Judith	Nurit	Ran	Ruthie	Jeff	Joseph	Meir T.
Break 11:40-12:00											
Unit 3	Theater	Suns, comets and stars	The wonderful world of mathematics	``Fantastic Beasts and Where to Find Them'' ecology	The inventions that changed the world	Mechanics and chain reactions	The chemistry of everyday	Tactical-strategic games	"Folded world"	The soul bird	The wonderful numbers
12:00-13:30	Neta	Meir P.	Gilad	discusses	Yvonne	Joseph	Michal	Meir T.	Esther	Judith	Liat

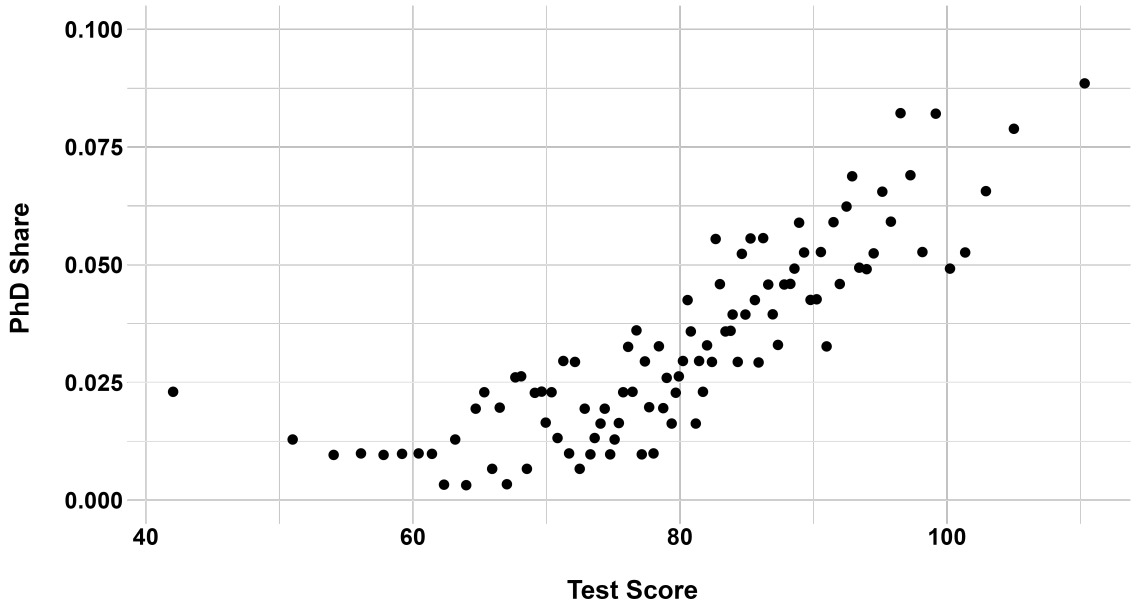
Notes: The figure presents the spring-term schedule for a third-grade pull-out class in one of the centers in our data, translated from Hebrew. Students attend the gifted center for one full school day and take three enrichment courses that do not duplicate the regular elementary-school curriculum. The schedule is illustrative; course offerings differ across centers and years.

B Data Appendix

Admission records. The historical admission files combine paper records from 1988–1995 with digital records from later cohorts. An approved government contractor digitized the paper files using optical character recognition. Approved CBS employees then compared the digitized records with the paper originals and corrected errors. CBS linked the records to the population registry using national identification numbers. The match was nearly complete, except for corrupted data in 1994.

School outcomes. The grade-5 and grade-8 examinations come from the national Growth and Effectiveness Measures for Schools program. Beginning in 2002, the program sampled

Figure A.2: Admission-Test Scores Predict Long-Run Outcomes



Notes: The figure shows the relationship between the second-grade admission-test score and the probability of obtaining a PhD by age 34 for applicants to the pull-out program. The horizontal axis reports the admission score, and the vertical axis reports the outcome indicated in the figure. This relationship is descriptive and is included to illustrate that the admission examination contains substantial information about later-life achievement; causal estimates rely on discontinuities at the admission cutoff rather than comparisons across the full score distribution.

schools for examinations in mathematics, Hebrew or Arabic, and English, with science and technology tested in grade 8. We standardize each test score by subject, grade, and year and then combine the standardized subject scores into the grade-specific outcome. High-school matriculation records cover the universe of examinations and include subject, level of study, credits, and scores. Separate annual student records report each student's school and homeroom assignment, which we use to identify gifted classes and second-grade classmates.

Higher education and labor-market records. The Council for Higher Education records enrollment of all Israeli students by institution, field, and year, as well as degrees completed in Israel and their completion year. The records also contain some recognized foreign degrees, although foreign degrees may be underreported when individuals have no reason to request formal recognition. Tax and employment records report annual wage income, self-employment labor income, and industry.

Family socioeconomic status. We observe parental education, parental income, family composition, parental country of birth, and school-level socioeconomic measures before the

admission examination. These variables are used for descriptive statistics, balance tests, covariate adjustment, and heterogeneity analysis.

Inferring admission thresholds. For each cell defined by program, gifted center, year, and gender, we choose the inferred admission cutoff that minimizes the number of recorded admission decisions inconsistent with score-based admission. We subtract this inferred cutoff from each applicant's score, so the running variable is the distance from the inferred cutoff in raw admission-score points, and pool the centered observations in the RDD analysis.

C Heterogeneity

The average effects mask important differences across groups. We focus on gender and family socioeconomic status because these dimensions are central to the debate about representation in gifted education.

C.1 Gender

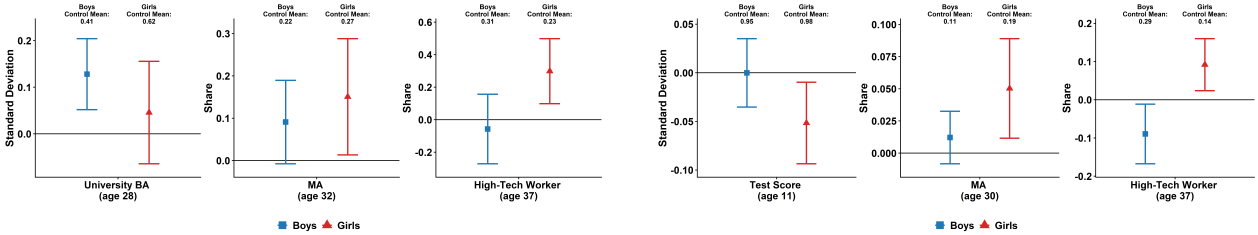
Girls are underrepresented in all three programs in the older cohorts. Figure A.3 shows gender-specific estimates for the three programs. Pull-out effects are particularly strong for women on master's-degree attainment and technology-sector employment; the full-time and afternoon estimates also show gender differences, although the pattern is less uniform across outcomes and programs.

These results distinguish representation from treatment effectiveness. The programs may admit too few girls while generating especially large benefits for those admitted. Policies that broaden identification and reduce barriers to female participation could therefore increase both equity and aggregate program benefits.

C.2 Socioeconomic Status

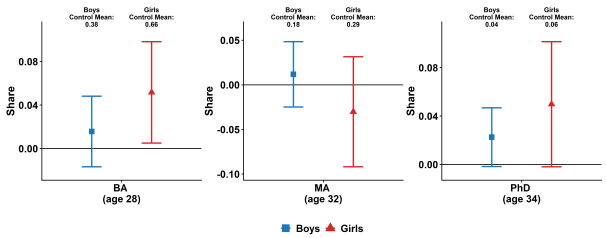
Students from high socioeconomic backgrounds are overrepresented among admitted students, but the effects are not uniformly larger for them. Figure A.4 presents estimates by family socioeconomic status. For the full-time and pull-out programs, heterogeneity is mixed

Figure A.3: Gender Heterogeneity by Program



(a) Full-time

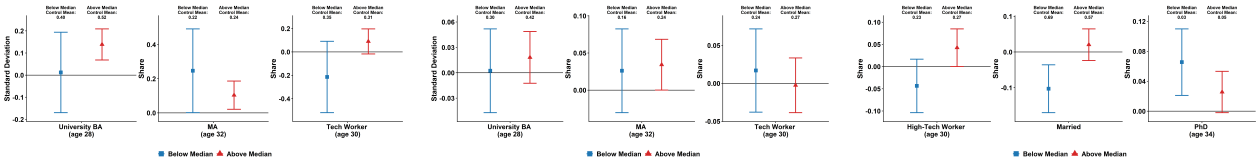
(b) Pull-out



(c) Afternoon

Notes: The panels report pooled RDD estimates separately for men and women. The pull-out outcomes are the grade-5 test score, master's-degree attainment by age 30, and technology-sector employment by age 37. The age-11 test score averages standardized scores in mathematics, Hebrew or Arabic, and English. The full-time panel presents university bachelor's-degree attainment, master's-degree attainment, and technology-sector employment; the afternoon panel presents bachelor's-, master's-, and doctoral-degree attainment. Markers are point estimates and vertical bars are 90 percent confidence intervals.

Figure A.4: Heterogeneity by Socioeconomic Status



(a) Full-time

(b) Pull-out

(c) Afternoon

Notes: The figure reports pooled RDD estimates separately for applicants classified as lower and higher socioeconomic status. Panels (a) and (b) show effects on university bachelor's-degree attainment, master's-degree attainment, and technology-sector employment for the full-time and pull-out programs. Panel (c) shows effects on doctoral-degree attainment, technology-sector employment, and marriage for the afternoon program.

across outcomes. In the afternoon program, the plotted technology-employment and marriage effects are smaller for lower-SES students, while the doctoral-education estimates show little difference across SES groups.