NBER WORKING PAPER SERIES

GIFTED CHILDREN PROGRAMS' SHORT AND LONG-TERM IMPACT: HIGHER EDUCATION, EARNINGS, AND THE KNOWLEDGE-ECONOMY

Victor Lavy
Yoav Goldstein

Working Paper 29779
http://www.nber.org/papers/w29779

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
February 2022

We thank the Central Bureau of Statistics for providing access to the data we use in this study in its protected research room in Jerusalem. We also thank Dr Anat Ben-Simon, general director of The National Institute for Testing, for helpful guidance and information about the University Psychometric Entrance Test. We thank Netanel Ben-Porath, James Fenske, Emma Duchini, Hessel Oosterbeek, and participants in seminars at Ben Gurion University, Bocconi University, CEMFI Madrid, Hebrew University, Pompeu Fabra, University of Warwick, and the CESifo Education Conference in Munich for valuable comments and suggestions. Lavy acknowledges financial support from the Israel Science Foundation, Falk Research Institute, and CAGE. Goldstein acknowledges the Azrieli Foundation for the support of an Azrieli Graduate Fellowship. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Gifted Children Programs’ Short and Long-Term Impact: Higher Education, Earnings, and the Knowledge-Economy
Victor Lavy and Yoav Goldstein
NBER Working Paper No. 29779
February 2022
JEL No. J01,J24

ABSTRACT

This paper examines the short-run and longer-term effects of gifted children’s programs (GCP). Using administrative data from Israel, we follow students who participated in a GCP, studying in separate gifted classes in high schools, and compare them to equally gifted students from localities where a GCP was not offered. Our results show that while GCP participation has tiny effects on high school academic achievement, it substantially influences university outcomes. This influence is manifested in the choice of field of study, a higher incidence of double majors, and an increased likelihood of pursuing advanced degrees. Interestingly, GCP participation does not affect earnings or employment in knowledge-based sectors, implying that gifted children do well in the labor market, regardless of participation in a GCP. Finally, participation in the GCP does not affect the likelihood of marriage or having children. Still, it positively affects the spouse’s “quality”, driven by marriages between GCP participants and their classmates. We discuss potential mechanisms by relating our findings to the literature in psychology about gifted children.

Victor Lavy
Department of Economics
University of Warwick
Coventry, CV4 7AL
United Kingdom
and Hebrew University of Jerusalem
and also NBER
v.lavy@warwick.ac.uk

Yoav Goldstein
Tel Aviv University
yoavg2@mail.tau.ac.il

1 Introduction

Human capital, especially at the high end of the ability distribution, is a crucial and significant factor for economic growth. The knowledge economy, including the high-tech sector, is considered the "workhorse" in the growth process in many developed countries. In Israel, this sector is regarded as the main driver of the national GDP growth in recent decades, contributing about a third of the national exports.[1] Gifted students are making a more significant part of the human workforce in these sectors and, therefore, receive special attention in many educational systems. However, despite the considerable amount of resources and time invested in this group, the evidence of their effect on enhancing employment and productivity in these sectors is quite limited.

This paper provides evidence of these issues by exploiting a long-existing gifted children's education program in Israel and unique administrative data that permits following gifted children from high school and into the labor market. We estimate the short-run and longer-term effects of gifted children programs (GCP) that started in high or middle schools in Israel. The program tracks the most talented students into gifted children classes, starting in 10th grade (or 7th grade in some schools). As a result, they receive more resources, a unique and accelerated curriculum, access to high-quality teachers, and attend university courses.

Based on administrative data, we follow twenty-two cohorts of GCP participants who graduated high school in 1992-2013. We use standardized test scores from different exams taken by students in Israel at different ages to select comparison groups of equally gifted students from other localities where GCP was not offered at the time. We also use pre-determined academic choices (as proxies for academic motivation) and family background measures to validate that the comparison students are similar to the GCP participants in these aspects.

We address the concern regarding systematic differences between the potential outcomes of gifted children in localities with a GCP and in other localities. First, we show that students in the localities with and without GCP have similar characteristics, including parental background and academic outcomes. Second, restricting the comparison group to include only students from larger localities without a GCP, or, alternatively, from localities with a GCP, does not change the results. Thus, unobserved differences between localities with and without a GCP are unlikely to bias our results. Finally, we also run a placebo exercise with students in regular classes in localities with a GCP (in different high schools). We use the same strategy to select a comparison group for these students and estimate the "impact" of studying in these regular classes. We find desired null results, further supporting the validity of our methodology.

[1] These data are from the Israeli Democratic Institute, Report on the future of the Israeli growth engine (available at https://www.idi.org.il/books/5370, retrieved on 27-02-2023).

We analyze the effects of GCP participation on outcomes ranging from high school to adulthood. We show that gifted children's academic achievements in high school are not significantly affected by GCP participation. We find minimal and mostly insignificant results when estimating the effects on matriculation test scores in different subjects. The effect on the mean composite score is zero. In the long term, we find no effects of GCP on the rate at which students gain undergraduate degrees, as almost all gifted children achieve this degree (above 90 percent). However, we do find impacts on their choice of field of study. Remarkably, GCP participants are much more likely to graduate with double majors than gifted students in the comparison group. Moreover, we find substantial positive effects on the attainment of advanced degrees.

Importantly, we also analyze how the GCP affects career outcomes. While all gifted children have much better outcomes than the average student, no difference exists between those who participated in a GCP and those who did not. Furthermore, we show that GCP participants and students in the comparison group have similar earnings and employment rates in the knowledge economy sectors. The lack of effects on labor market outcomes persists until advanced career stages. These findings suggest that gifted students have successful careers, but the GCP has no significant contribution to their career paths.

When examining the personal outcomes, we find that the GCP does not influence the likelihood of marriage or having children. However, it positively impacts the "quality" of spouses, measured, for example, by their test scores, which is attributed to the higher likelihood of GCP participants marrying their classmates. Furthermore, beyond marital considerations, it is observed that GCP participants tend to work alongside other GCP participants, reinforcing the notion that the GCP influences the social connections of gifted students. These findings may suggest that a significant benefit of the GCP for gifted students is the exposure to similarly high-achieving peers.

In the short-term, medium-run, and adulthood, these comprehensive results are not significantly different for different groups of gifted children who participated in GCP. We examine how the effects vary by gender, socio-economic status (SES), giftedness level (or academic ability), and length of participation in GCP. We do not find significant heterogeneity in any of these dimensions.

To ensure the validity of our findings, we conducted extensive robustness checks. These were designed to test the causal interpretations of our results. Initially, we addressed concerns about reliance on specific model specifications. Our findings remained consistent across various alternative specifications, including those utilizing advanced machine-learning techniques for estimating propensity scores. Additionally, we applied a complementary approach that focused exclusively on the highest-performing individuals within the ability distribution. This comparison between top achievers in GCP and their counterparts in other regions is complementary because it builds a comparison group of gifted students in a less prescriptive manner. This analysis yields results in line with our main strategy, further strengthening the validity of our conclusions.

In addition, we validate our results' robustness by utilizing various samples and variables. We employed three distinct sets of test scores, each administered at different ages, as proxies for ability. Available to various cohorts, these tests offer a comprehensive view of the abilities assessed. The uniformity of our findings across these diverse samples and age-specific test scores not only strengthens their reliability but also highlights the consistent results across different periods.

The evidence we present in this paper contributes to the few recent studies on the causal effect of GCPs. Card and Giuliano (2014) apply a fuzzy regression discontinuity (RD) design to estimate a GCP's impact on primary school test scores. The GCP in focus brings together gifted students in classrooms with other high achievers and offers an enriched curriculum. This study finds no test score improvements.[2] Bui et al. (2014) also examine the effect of GCP on test scores in the South-western U.S. Using either a fuzzy RD design comparing students scoring just above or below the GCP admission cut-off or exploiting a lottery in oversubscribed middle schools offering the GCP program, the authors find no significant positive effect on student performance. While these studies focus on separate gifted classes in primary and middle schools, our research extends this investigation to high schools, demonstrating that attending such classes at this level also does not improve test scores.

In a recent study, Cohodes (2020) examined the short and long-term effects of participation in dedicated classes for high achievers in grades 4 to 6 in Boston public schools. While she also did not find statistically significant effects on test scores, the results provide evidence that participation in these specialized classes enhances college enrollment, particularly for minority students. In our context, baseline higher education enrollment rates are already relatively high, and as a result, we did not observe enhanced enrollment. This discrepancy may arise from the different populations participating in the programs. In Cohodes (2020), the program is formally described as for high-achievers rather than specifically for the gifted. However, we did find enhanced higher education outcomes in our setting, particularly regarding the chosen field of study, double majors, and pursuit of advanced degrees.

Other studies have also delved into the effects of different types of gifted children's education programs. Redding and Grissom (2021) find that participation in a gifted enrichment program in public primary schools is associated with modest achievement gains. Booij et al. (2016) examine the impact of a gifted secondary education program,

[2] However, it has been demonstrated that participation in these classes yields significant and positive effects on the achievement of high-achieving non-gifted students (Card and Giuliano, 2016).

which is an individualized pull-out program in a specific school in the Netherlands. Like earlier studies, they employ a fuzzy RD design to estimate the impact on those at the program's margin of acceptance. They find that participants achieve higher grades, express stronger beliefs about their abilities, and make more advantageous university field of study choices regarding associated financial returns. In a separate paper, they analyze similar programs in other schools to investigate the effects on academic performance Booij et al. (2017). They employ different strategies and demonstrate that the effects are more pronounced for students farther from the admission cutoff.

This paper presents several significant contributions to the existing literature. Firstly, it is based on an experienced gifted children's program that has been active for over three decades. This extensive experience allows for a comprehensive assessment of long-term outcomes. Specifically, the study examines critical university choices, including fields of study and attaining advanced degrees, outcomes particularly relevant to gifted children. Additionally, it tracks GCP participants beyond their degrees, investigating their employment and family formation patterns. The paper also analyzes GCP participants' impact on the knowledge economy, examining their integration into high-tech sectors and academic institutions. Secondly, an additional important contribution of this paper lies in analyzing treatment heterogeneity based on giftedness levels. This differentiates it from earlier studies that predominantly used RD designs, focusing mainly on marginally eligible students for such programs. Furthermore, the paper makes a significant distinction in estimating the treatment effect based on the duration of GCP participation.

We relate some of our findings to theories and hypotheses in the psychological literature regarding gifted children. It includes studies about the affective and personality development of gifted children. The literature on "big fish small pond", which suggests that students may feel less competent in more competitive environments, is perhaps key in understanding our finding that GCP has no effect on test scores in high school (see Marsh et al., 2008, for a review). Of particular relevance to us are also studies on the effect of labeling (being part of a gifted program) and excessive parental expectations and pressure from teachers and social networks (e.g., Robinson et al., 2002; Pfeiffer et al., 2003). Related literature coined the term 'the gifted paradox.' Gifted children have an ability that can be used for a meaningful process of self-exploration to form an identity. Still, external pressures curtail this process and lead them to choose, for example, prestigious professions. This tends to hasten the process of identity formation and limit self-exploration. This paradox is related to the term "multipotentiality", which characterizes gifted children in GCP (Leung et al., 1994; Kerr and Colangelo, 1988). Our findings that GCP causally directs gifted adolescents to double majors at the university are likely related to this paradox.

The paper is organized as follows. The following section (2) describes the gifted education programs in Israel and elsewhere. Section 3 presents the data, and section 4 presents the empirical methodology. Next, we present the results in section 5. We validate their robustness and persistence in section 6, and examine the heterogeneity of the GCP effects in section 7. Finally, section 8 provides conclusions and further discussion.

2 Context and Background

2.1 Gifted Children Programs

In most countries, fostering gifted students' talent is essential in the knowledge economy and crucial for securing new generations of scientists, creators, and innovators. Yet, how to deliver gifted education is at the center of a longstanding and still hotly debated topic in education policy circles. In many countries, introducing specific practices for talented children dates back to the 1960s (Vrignaud et al., 2005; Mönks et al., 2005; Boettger and Reid, 2015). Over time, these included interventions targeted at different age groups, from early enrolment in primary school to grade skipping, curriculum enrichment, extracurricular syllabus, and summer camps. Remarkably, despite this longstanding debate, there is little causal evidence on the relative effectiveness of gifted education programs for different targeted groups and outcomes.

Different countries and school districts in the same country also adopt different selection procedures. Early GCPs used intelligence assessment (e.g., I.Q. scores) as the basis for eligibility. Still, this selection method has been strongly criticized as I.Q. tests are argued to be ethnically or racially biased. As an alternative, researchers and practitioners have suggested that eligibility should be based on a combination of cognitive and non-cognitive measures.

2.2 Gifted Children's Education in Israel

By the late 1980s, Israel had developed a separate study program for highly gifted students throughout grades 3–12.[3] This program incorporated elements of enrichment, extension, and acceleration. Parallel to this, some universities started to offer education and training to teachers of gifted children. By 1994, the Ministry's Department for Gifted Education had extensive responsibilities, including testing children in some cities, establishing unique enrichment frameworks, and instructing teachers and field workers.

[3] The material presented in this section draws details from https://giftedphoenix.wordpress.com/2012/11/15/gifted-education-in-israel-part-one (retrieved on 06-09-2021).

Since then, three types of GCP have been offered: (1) A weekly program organized by a city or school district, often starting in third grade and continuing until the end of primary school (6th grade), including weekly enrichment days in pull-out sessions. (2) Special classes in one of the regular city schools enable gifted students to be taught in separate classes.[4] The learning content is based on the standard school curriculum. Still, it incorporates advanced concepts and topics, various teaching methods, and joint teaching with university staff. (3) An afternoon enrichment program.

Finally, a 2004 reform consolidated the country's GCP into a national program to develop Israeli gifted education. It embraced the two-morning frameworks–weekly enrichment days and special school classes. As a result, the number of special classes operating in secondary schools has expanded (from 11 to over 20). Additionally, specific localities started offering GCP programs for middle-school students during these years.

This paper focuses on upper secondary gifted children programs (type 2 above) because they are numerous, offer a meaningful sample size for analysis, and resemble many of the GCPs in Europe and the U.S., offering more external validity to this paper's findings.[5] Admission to these programs is based on an intelligence test undertaken during the year preceding the program. During the 1990s, there were gifted classes in 11 high schools in 10 localities in Israel, most in the major cities.

Throughout the paper, we analyze the outcomes of GCP participants in these eleven oldest programs who graduated high school between 1992 and 2013. In the primary analysis, we restrict our attention to those who graduated between 2006 and 2010, as we observe their pre-determined test scores and labor-market outcomes at relevant ages. About half of the students in this sample also participated in the GCP during middle school. We use this variation to estimate how the effects of a GCP vary by the age students started the program, namely length of exposure. Furthermore, we analyze the outcomes of earlier and later cohorts as supplementary analyses, enabling us to validate our results' robustness and persistence until late ages (up to 46).

2.3 Israel's High School and Higher Education Systems

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 (or "Bagrut") if they pass a series of national exams in core and elective subjects between 10th and 12th grade. Depending on difficulty, students choose to be tested at various proficiency levels, each awarding one to five credit units per subject. Advanced-

[4] One exception is a residential school for the gifted that serves children from all over the country and is located in Jerusalem.

[5] Note that we focus on middle and high schools since dedicated classes for gifted students were very rare in primary schools in Israel during our research years.

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 matriculation certificate. Courses that award five credits are equivalent to Advanced Placement courses in the US high school system.

Matriculation is a prerequisite for university admission, and receiving it is an economically important educational milestone. About 52% of all high school seniors received a matriculation certificate in the 1999 and 2000 cohorts (Israel Ministry of Education, 2001). The rates among gifted children are much higher (more than 90% among our sample's gifted students). Furthermore, a typical study program for gifted children includes several subjects at an advanced level (where the minimum requirement is only one). A study program that includes several subjects at an advanced level is challenging and demanding, and only very talented or gifted children follow it. For more details on the Israeli high school system, see e.g., Lavy (2020).

Israel's higher education system includes ten universities (one confers only graduate degrees) and 50 colleges that confer undergraduate degrees (some also give master's degrees). All universities require a matriculation certificate for enrolment. Most academic colleges also require a matriculation certificate, though some look at specific components without requiring full certification. It is typically more difficult for a given field of study to be admitted to a university than a college.

3 Data

We use an administrative database from Israel's Central Bureau of Statistics (CBS), available at their protected research room in Jerusalem. The data is based on merged datasets from multiple sources such as the population registry, the Ministry of Education (information on primary, middle, and secondary education), the Higher Council of Education (post-secondary education), and the Israel Tax Authority (information on earnings and employment). For more details on the database and its sources, see Appendix A. The baseline sample includes information on all individuals in Israel who were born between 1970 and 1995.

3.1 The Analysis Samples

We do not observe the gifted screening exam scores. Moreover, since screening exams for gifted children were administered primarily in cities with GCP, no such systematic test scores are available for selecting a comparison group from other localities. We, therefore, opt for other ability measures. We define different samples according to the ability measures we observe for these students. Our data includes two different kinds of exams that measure ability and intelligence. The first is the national Meitzav exams taken in four subjects (science, math, English, Hebrew) during primary school (5th grade) and middle school (8th grade). The second is the University Psychometric Entrance Test (UPET), which includes test scores in three domains (quantitative, verbal, and English).[6]

The clear advantage of using the Meitzav test scores is that their timing is before participating in a GCP. However, the limitation is that these national exams were introduced in 2002, allowing us to observe 8th-grade (5th-grade) test scores only for high school graduates of 2006 and later (2009 and later). Thus, in our main analysis, we use the sample of 2006–2010 graduates with their 8th-grade test scores as the proxy for ability. In 2020, the latest year in the labor market data we use, the youngest cohort in the main sample is 28, while the oldest is 32. This age range ensures that most individuals have completed their undergraduate degrees and usually are well integrated into the labor market. The sample includes all students in these cohorts who participated in the 8th-grade tests, about half of the students in each cohort. Before matching, it includes 626 GCP participants and 63,644 students from other localities, which are included in our comparison group pool (from which we identify equally gifted children, as described in section 4).

We also analyze the outcomes of two additional samples with different limitations. First is the 2009–2013 graduate sample, for which we observe 5th-grade test scores. The limitation of this sample is that we observe their labor-market outcomes only until ages 25–29. Second is the 1992–2005 graduate sample, for which we observe labor-market outcomes at ages 33–46. The limitation of this sample is that the only test scores we observe are the UPET scores. However, we show evidence that using UPET scores as the ability measure for identifying gifted children from other localities is valid.

3.2 Definitions of Outcomes

Class characteristics. We use class characteristics as outcome variables to demonstrate how gifted classes differ from regular classes. This includes the number of students, the share of females, average peers' academic outcomes and background characteristics, and the share of students taking university courses during high school. Unfortunately, we do not observe teachers' characteristics, which may also be better in gifted classes relative to regular ones.

High school achievement. We use matriculation test scores in mandatory subjects as the outcome variables and calculate the mean composite matriculation score for each individual in our sample.

[6] The UPET is required for university applicants in Israel and is administered by The National Institute for Test and Evaluation (NITE). According to NITE, the UPET is a tool for predicting academic success at higher education institutions in Israel.

Higher education. First, we define an indicator for getting an undergraduate degree. To study how GCP affects decisions regarding the field of study in university, we create dummy indicators for areas of study that lead to employment in the knowledge economy, specifically STEM, and its components: math and computer science, engineering, physical sciences, and biological sciences.[7] We follow the grouping definition of the CBS. We also define indicator variables for achieving advanced degrees.

Employment. We use indicators for employment (a non-zero number of months of work in a given year and a non-zero income) and self-employment (non-zero business income). We also define indicators for employment in the knowledge economy sectors. Using a three-digit sector code, we focus on the following sectors. *High-tech Manufacturing industries*: Pharmaceutical products for human and veterinary uses; Office and accounting machinery and computers, electronic components, electronic communication equipment; Aircraft. *High-tech Services industries*: Telecommunications; Computer and related services; Data Analysis; Research and development. *Academic*: Colleges; Universities. *Knowledge*: any of the above.[8]

Earnings. We focus on the total annual earnings. We use the earnings rank conditional on age as the main outcome variable. We also analyze the natural log of the earnings and nominal earnings.[9] The exact earnings data is also available for our sample's parents for the same years.

Personal outcomes. We use indicators for marriage and having at least one child. Having the same data for marital partners, we measure their "quality" based on GCP participation and test scores.

4 Methodology: Identification of GCP Short and Long-Term Effects

Previous studies used fuzzy RD designs to estimate the effect of GCP programs in the U.S. (Card and Giuliano, 2014; Bui et al., 2014). This design exploits the admission cut-off to GCP. It yields a local average treatment effect of providing gifted education services to students on the margin of gifted child qualification. However, this paper uses an alternative identification strategy to understand how GCP affects the achievement of gifted children beyond the marginal student. We chose gifted children from cities

[7] About half of the STEM graduates work in knowledge-producing sectors relative to less than 10% in the general population.

[8] We further validate the reliability of the labor market outcomes by comparing their means to the respective statistics based on labor survey data available for a sub-sample of individuals in our sample. However, we do not use these data in our analysis because the sample is small.

[9] We exclude observations that deviate by six or more standard deviations from the mean to account for earnings outliers. Only a few observations are dropped; this procedure does not affect the results.

where GCP was not offered at the time as a comparison group.

4.1 Identifying Gifted Students from Other Cities

Our empirical strategy raises the challenge of selecting gifted students from other cities. Specifically, to provide a valid counterfactual, the comparison group's students should be identical to GCP participants in any relevant aspect (that affects the selection into GCPs). Thus, we use three sets of variables to select the comparison group, each capturing different characteristics that might affect enrolment in a GCP. The first important group of characteristics is intelligence and academic ability, perhaps the most important factor in the selection process. We use test scores that measure general intelligence and ability. In our main analysis, we rely on the scores of the national Meitzav exams taken in four subjects (science, math, English, Hebrew) during 8th grade. We standardize all test scores at the cohort level.

Individuals are selected (or self-selected) for GCPs not just by their intelligence and academic ability but also according to their academic motivation and aspirations. While we do not observe any survey data that measure self-reported motivation, we observe academic choices that reflect academic motivation. We use the high school study program, which is individually chosen before the start of 10th grade and most likely reflects the student's academic motivation and ambition at this stage. Since a student's study program is probably pre-determined, we can use it to match GCP participants to students in the comparison group.

Additionally, individuals might be self-selected for GCPs according to their socioeconomic status and family backgrounds. Thus, we also used measures of parental education and country of birth, number of siblings, and birth order when choosing the students who made up the comparison group.

4.2 Constructing the Comparison Group

We construct a comparison group using the following two-step propensity score matching algorithm.[10] We start by estimating the propensity score equation using the sample of GCP participants and students from other localities (the comparison group pool):

$$P_1(X_i) = Pr(GCP_i = 1|X_i) \tag{1}$$

[10] Rosenbaum and Rubin (1983) proposed an approach that circumvents the curse of dimensionality when using selection on observables to identify causal effects. They prove that if treatment assignment can be ignored given x, then it can be ignored given any balancing score that is a function of x, particularly the propensity score. See Abadie and Cattaneo (2018) for an updated survey of econometric methods for program evaluation and a useful comparison of matching/propensity score models with other methods.

Where GCP is an indicator of participation in a GCP. We use a Logit specification in our main analysis, but we also validate that our results are robust for non-parametric estimation using a gradient boosting algorithm (Chen and Guestrin, 2016).

X_i is a vector of individual covariates. We include the standardized 8th-grade test scores in each domain (with quadratic terms) as measures of ability. As proxies for academic motivation, we include indicators for achieving five matriculation credits in English and math (mandatory subjects, but the minimum number of credits is three) and in the fifteen most common elective subjects among GCP participants relative to the general population. We also include an indicator for a high (above median) number of total matriculation credits. Finally, we include the following family background variables: The father's and mother's years of schooling (with dummies for each level), an indicator for having at least two siblings (the median in our sample), an indicator for being the oldest sibling in the family, and an indicator for individuals whose both parents were born in Israel. We also include cohort indicators.

Then, we match GCP participants to the comparison group using the nearest neighbor without replacement. We include in our sample only matches in a caliper of 0.1 standard deviations of the propensity score and with the same sex, the same religious status of the school, and the same matriculation track (regular or technological). We also use alternative specifications to validate that our results are robust.

Figure 1 presents the propensity score distribution before and after the matching. We matched 555 of the sample of 626 GCP participants. The unmatched are mainly from the top of the propensity score distribution.[11] The propensity scores of the GCP participants and their matched counterparts are perfectly aligned and not distinguishable after the matching. Figure 2 presents the distributions of the 8th-grade scores before and after the matching. As expected, there are substantial differences between GCP participants and the sample of students from other localities before matching since most students in these localities are not gifted. However, the matching eliminates most of these differences in test scores, as all the distributions become statistically indistinguishable.

Table 1 and Table 2 show a detailed summary of descriptive statistics for our sample, including a test for mean differences between the pre-determined outcomes of GCP participants and the matched comparison students. Columns 1 and 2 show the averages of the comparison and treatment (GCP) groups. Column 3 shows the estimated difference. Table 1 shows that both groups are perfectly balanced regarding parental characteristics. Interestingly, the groups are balanced on variables not included in the matching specification, such as parental earnings, parental country of birth, and pater-

[11] This is reflected by their better characteristics. For example, the unmatched students have higher 8th-grade test scores by 0.15-0.3SD relative to matched GCP participants. Thus, it might be harder to find a match for them.

nal age at birth. For example, the father's (mother's) average yearly earnings in 2003 were 153,000 (85,000) NIS in the treated group and 165,000 (78,000) in the comparison group. These differences are not statistically different from zero. This suggests that unobservable variables are also likely to be balanced, given the balance of observed variables that were not specifically matched.

The evidence in Table 1 shows that GCP participants and the comparison group's students come from higher socioeconomic backgrounds than regular students. For example, mean mother and father years of schooling are around 15 years for these two groups, higher than among non-gifted children (where parental years of education are about 14).

Table 2 shows that both groups are balanced regarding student characteristics, with all estimates being statistically insignificant. Unsurprisingly, gifted children study in extensive matriculation programs with around 29 credits on average (relative to the minimum of 20). They also participate in extensive math and English studies at very high rates (about 75% and 90%). We also find balance in the likelihood of studying any of the scientific elective subjects, which are also the most common electives among gifted students–computer science, physics, chemistry, and biology.

4.3 GCP Effects Estimation

We estimate the following controlled regression using our matched sample of GCP participants and equally gifted children from other cities:

$$Y_i = \alpha + X_i'\beta + \tau \times GCP_i + \varepsilon_i \tag{2}$$

Where τ is the coefficient of interest, capturing the effect of GCP participation on the outcome. The standard errors of the program effects estimates were clustered at the school level.[12]

This propensity score matching and regression combination allows for enhanced robustness to misspecification. As long as the parametric model for either the propensity score or the regression functions is specified correctly, the resulting estimator for the average treatment effect is consistent. This notion is termed "double robustness", which is discussed in Robins and Ritov (1997); Imbens (2004).[13]

[12] We also validate that the standard errors are similar to clustered bootstrapped standard errors or to the correction offered by Abadie and Imbens (2008) (see Appendix Table A1).

[13] See Abadie and Imbens (2002) for details regarding using OLS with the matching procedure weighting. We note that OLS with controls will estimate an average effect for the whole population, which is inappropriate in our context given that only gifted children can be treated, namely, participate in GCP. The propensity score estimate is the average effect on the treated, which is our parameter of interest.

4.4 Concerns About Causal Identification

Our identification strategy relies on a conditional independence assumption (CIA) and on the quasi-natural experiment that GCP is not offered in many cities in Israel. Here, we discuss three potential concerns regarding our empirical strategy. The first potential concern is that families may relocate based on access to the GCP program in a given locality. To address this, we examined whether families with GCP participants exhibited a higher mobility rate before 10th grade compared to families with gifted children who did not participate in a GCP. However, our findings indicate no such differential mobility rate. In fact, the rate of families relocating between the 6th and 11th grades is 23.3% in the comparison group and 20.5% in the treatment group. The difference is negative at -2.5 percentage points and insignificant (with a standard error of 2.8), thus mitigating concerns about increased relocation among families of GCP participants.

The second potential concern is that there could be systematic differences between the potential outcomes of gifted children in cities with a GCP, typically larger towns, and those of gifted children in other cities. For example, the educational and economic preferences might differ between individuals in large and smaller towns. First, we show that although cities with GCP are larger, students in these cities have characteristics similar to those in other cities, including parental background and academic outcomes (see Appendix Table A2). Second, restricting the comparison group to include students from larger cities yields the same results. Third, we show that the results are similar when we use a comparison group of students from the same cities where GCP is offered. Thus, it is unlikely that this channel would bias our results.

The third potential concern is missing an important variable from the matching specification. This should be a variable that is important in the selection process for GCPs and also affects the outcomes that we analyze. Given that our data set includes detailed information on each student's academic ability, motivation, and family background, we think it is unlikely such a variable could bias our results. This is further supported by the observed balance between both groups of gifted children after the matching, which extends to variables not initially included, such as parental income.

Moreover, we also provide suggestive evidence that supports the causal interpretation of our findings by running a placebo exercise to study the "effect" of studying in regular classrooms. First, we define a new treatment group that includes non-gifted children who study in regular classrooms in other high schools in localities with a GCP. Then, we implement an identical matching algorithm to select a comparison group for these students and estimate the conditional difference in primary outcomes between both groups of students. As discussed in Section 6, this exercise yields desired null results concerning the primary outcomes, further supporting our analysis's validity.

5 Short-, Medium- and Long-Term Effects of GCP

In this section, we present the estimation results on the effects of GCP participation. We start by analyzing the primary sample, including GCP participants, and matched comparison students who graduated high school between 2006 and 2010. In Section 6, we extend the analysis to additional samples to show that the results are robust and persistent, and in Section 7, we analyze the heterogeneity of the effects.

5.1 Class Environment

We begin by examining the impact of GCP participation on the classroom environment. Table 3 compares the class-level characteristics of gifted students in GCPs with those in the comparison group. Notably, GCP classes tend to have smaller sizes, with an average of nearly seven fewer students. GCPs also exhibit a higher proportion of male students than regular classes, with a difference of 9p.p. Furthermore, students in GCPs are exposed to peers with higher socio-economic backgrounds, as indicated by their parents' years of schooling, and with better academic achievements, as evidenced by their higher UPET scores and greater likelihood of attaining academic degrees.

In addition to the outcomes presented in the table, institutional background information indicates that GCP teachers typically have higher qualifications and receive additional training. Unfortunately, our dataset does not provide direct information about the teachers. Furthermore, GCPs often provide access to more advanced curricula; in some instances, students may even have the opportunity to take university-level courses.

5.2 High School Achievement

Participants in GCP attend smaller classes with more resources and higher-quality peers. How does this affect their high-school achievement? Figure 3 compares the mean composite matriculation score distributions of GCP participants and matched comparison students. The distributions are statistically indistinguishable, the averages are also almost identical, and the difference between them is small, negative, and insignificant. This implies that the GCP does not enhance matriculation achievement on average. Additionally, Appendix Table A3 shows the estimates for the average GCP effects on test scores in all compulsory subjects. Most estimates are small and insignificant.[14]

[14] Note that we report only the estimates based on Equation 2, including all control variables throughout the paper. Still, we also validate that the results are similar if we exclude them. This is not a surprise;

The pattern of no test score gains for GCP participants in matriculation exams is especially intriguing given the abundance of educational inputs that GCP participants enjoy relative to the comparison group we use. Although the impact of smaller class sizes on outcomes remains uncertain, with mixed evidence in the literature (see, e.g., Angrist et al., 2019), one might anticipate that the presence of more capable peers, high-quality teachers, and advanced curriculum would have an effect. So, what can explain the lack of positive effect of GCP on achievements at the end of high school exit exams?

First, as mentioned earlier, the GCP's studies program incorporates advanced concepts and topics that are not directly relevant to the matriculation exam material. It often emphasizes and encourages learning outside the standard curriculum. Additionally, gifted students' matriculation test scores are typically very high even without enrolling at a GCP, allowing them to enter most university degrees. Thus, it is very plausible that GCP participants get other educational benefits not manifested in higher test scores.

Alternatively, the minor effects on high school achievement could be due to the potentially adverse psychological effects of the change in within-class ordinal ranking regarding ability. When academically gifted students are placed in self-contained programs, they usually experience a new environment with equally competent peers, more challenging materials, and more rigorous requirements. One reality they inevitably must encounter is a more talented peer group than they are used to in a regular classroom. This can be harmful because individuals, particularly those who might already feel insecure, are likely to think that the very talented people have touted them.[15] They may also find that the top student status they have enjoyed in the regular classroom is no longer a sure thing, as there are potentially more talented people in the new peer group.

Therefore, when two students of the same ability or achievement level are placed in different classrooms or programs, the one with the high-ability group tends to temporarily lower self-concept in respective domains than those with the less able peers. This effect has been labeled the Big Fish Little Pond Effect or "BFLPE" (Marsh and Parker, 1984; Herrmann et al., 2016).[16] Although the BFLPE model is not specific to

we also show that the groups are balanced in any important measure (in subsection 4.4).

15 Theoretically, this could also be beneficial because a peer group of equal academic caliber gives personal validation to one's identity and mutually reinforces each other's talents and interests. But, the literature on GCPs emphasizes the negative impact.

16 Recent papers in the economics of education have documented this mechanism in other contexts. Elsner and Isphording (2017) show that students' ordinal rank significantly affects educational outcomes later in life, such as finishing high school, attending college, and completing a 4-year college degree. Exploring potential channels, these authors find that students with a higher rank have higher expectations about their future careers, a higher perceived intelligence, and receive more support from their teachers. Murphy and Weinhardt (2020) show that ordinal academic rank during primary school impacts secondary

gifted programs, facets of the BFLPE have been examined with gifted and high-ability students from the early elementary years (Tymms, 2001) to the college years (Rinn, 2007). The practical implications are obvious and have already produced repercussions in the gifted education community (e.g., Plucker et al., 2004; Dai and Rinn, 2008).

In our context, GCP participants moved from an environment in middle school where they were most likely at the very top of the ranking in their class to a class with peers who were, on average, equal. As a result, their rank order most likely declined. Earlier studies from Israel have shown that gifted students who move from heterogeneous classes to homogeneous classrooms where all students are gifted are also subject to BFLPE. Studies have shown that this change lowers their academic self-concept and increases their anxiety (Marsh and Parker, 1984; Zeidner and Schleyer, 1999; Marsh and Craven, 2002; Preckel et al., 2008).

Lastly, it is worth considering the potential impact of gender composition in gifted classes on student achievement. The proportion of male students is 9p.p. higher in GCPs compared to the comparison group's classes (see Table 3). Research has demonstrated that a higher presence of females in a class can have positive effects on students' achievement (e.g., Lavy and Schlosser, 2011). Therefore, the lower representation of females in GCPs might have adverse effects.

5.3 Higher Education Outcomes

We continue our analysis by investigating the impact of GCP participation on higher education outcomes, as presented in Table 4. Initially, we observe that GCP participants have a significantly higher enrollment rate in university courses during high school than their gifted peers in the comparison group (first row of Panel A). This phenomenon may be attributed to the fact that many GCPs offer students the opportunity to begin university studies during high school and provide support throughout this process.[17] Consequently, GCP participants also complete their undergraduate degrees at younger ages. For instance, the probability of attaining an undergraduate degree by 25 is 23.8% for GCP participants, compared to 14.8% for students in the comparison group. This 9-percentage-point difference is statistically significant at the 99% confidence level or higher, with a standard error of 2.4p.p.

school achievement independent of underlying ability. In addition, they found significant effects on test scores, confidence, and subject choice during secondary school, even though they had a new set of peers and teachers unaware of their previous ranking in primary school.

17 It's worth noting that one potential concern in interpreting this result causally is the possibility that our comparison group students reside in localities that are farther from universities. However, it is important to highlight that the results remain consistent even when we use a comparison group of students from the same localities, indicating that proximity to universities is not the driving factor behind these outcomes.

However, this difference becomes statistically insignificant when considering indicators for enrolling in and completing an undergraduate degree at any age, as nearly all gifted students pursue a degree. We also examined the likelihood of achieving a degree from an elite university and found an insignificant increase of 3.4 percentage points, with a standard error of 2.8 percentage points.[18]

Panels B and C of the table reveal that GCP participants pursue advanced degrees to a greater extent, with substantial effects. The increase in MA degree enrollment is approximately 10p.p. (about a 40% increase, statistically significant at the 99% level),[19] and the rise in PhD enrollment is around 3p.p. (almost a 90% increase, statistically significant at the 95% level). The boost in advanced degree attainment is also statistically significant and large, although it is somewhat less precise due to ongoing studies among some students.[20] Finally, as is shown in Panel D, we do not find a statistically significant change in the likelihood of pursuing medical degrees.

Table 5 shows the estimated GCP effects on university fields of study. Panel A focuses on STEM fields, including math, computer science, engineering, physical sciences, and biological sciences. While there is no effect on the likelihood of achieving any STEM degree, there is a significant composition change, a movement from engineering programs to computer science programs. The estimated effects are very large. The likelihood of studying engineering declines by 6.0p.p, amounting to a fall of 26 percent. The likelihood of studying computer science increases by 4.8p.p, implying a 24 percent rise.[21]

The program substantially impacts the likelihood of attaining a double major. As shown in Panel B, the estimated effect is statistically significant and substantial, amounting to 5.8p.p. relative to a baseline of 18.7%. This implies a 31% relative increase in the likelihood of pursuing a double major. Additionally, the results in this panel reveal that the increase in double non-STEM majors is statistically significant

[18] According to the Israeli CBS, three universities are elite. These are Tel Aviv University, the Hebrew University, and the Technion.

[19] Note that this result is highly statistically significant, even when applying a conservative Bonferroni correction for multiple hypothesis testing. The standard p-value for the increase in testing is 0.0002. Therefore, even if we take a very conservative approach of calculating the Bonferroni-corrected p-value for nine tests, we get a significant result ($p = 0.0022$).

[20] In 2020, the final year of our data, approximately 28% of the students in our matched sample are still enrolled in degree programs. However, we verify that this rate is balanced between GCP participants and the comparison group, and most of these students are pursuing advanced degrees and have already completed their undergraduate studies. Importantly, they are already well-integrated into the labor market.

[21] When analyzing non-STEM fields, we find a marginally significant increase in the likelihood of having any non-STEM major. This increase is by 4.7p.p. from the baseline rate of 30%. When analyzing non-STEM fields separately, we find a large increase in the likelihood of studying the humanities (3.6p.p. increase from the baseline rate of 4.8%). This increase is significant at the 95% level ($p = 0.02$). The likelihood of studying in other non-STEM fields is not changing.

and of considerable magnitude (42%). While the increase in double STEM majors is statistically insignificant, it is still noteworthy in terms of magnitude (30%).

Graduating with a double major can be related to the multipotentiality of gifted children. This concept is defined as "the ability to select and develop any number of career options because of a wide variety of interests, aptitudes, and abilities" (Kerr and Erb, 1991, p.1). Multipotentiality is widely cited as a characteristic of the most gifted individuals with the ability and interest to pursue various activities and goals, especially related to career choice (Sajjadi et al., 2001; Sampson Jr and Chason, 2008). This effect may be activated and enhanced in an environment where giftedness status is formally recognized as in a GCP environment.

The evidence shows how significant GCPs' effect shapes adolescents' university choices. Realizing academic potential is often perceived as acquiring higher education, impressive academic achievements, or pursuing a prestigious profession. But what motivates gifted adolescents to make future professional choices and the themes that guide them? To what extent does the environment impact these choices? Studies in educational psychology on forming gifted adolescents' identity (Zeidner et al., 2005; Zeidner and Shani-Zinovich, 2015) provide insights into these relevant questions for understanding and interpreting our results. They argue that the desire to realize their potential and the concern not to choose areas considered "potential waste" is a central theme among gifted adolescents, especially those enrolled in gifted classes. The label "gifted" impacts their choices; they are affected by their expectations to make the most of their high abilities, i.e., their potential, and exhibit a future focus that does not characterize non-gifted adolescents. They feel obligated to realize their potential in its conventional sense. This leads to an interesting paradox–precisely, those with high abilities who can choose any field of study are those who feel that they have only a limited range of options. In their experience, they are limited to the same possibilities that will be considered to realize their potential.

5.4 Labor Market Outcomes

An important question regarding gifted children's education is whether participation in a GCP significantly affects their career outcomes. We provide the first evidence regarding this issue by studying the long-term effects of GCP participation on early career outcomes (ages 28–32). Figure 4 compares the earnings distributions of GCP participants and matched comparison students. Both groups of gifted students earn more than non-gifted students, but there are no differences in earnings between GCP participants and the matched comparison students. Panel A of Table 6 shows estimates for the average effects on earnings, their natural log, and their rank (conditional on age). We do not find any evidence that GCP participation affects these outcomes. Consistent with the impression of Figure 4, we also find that the likelihood of becoming a top 10% earner does not change significantly. Additionally, Panel B of Table 6 shows that the GCP has no significant effect on the likelihood of being employed or self-employed.[22]

We also examine whether the GCP directs more talent to the economy's knowledge-producing sectors. About forty percent of gifted children were employed in 2020 in these sectors, including high-tech services and manufacturing and the academic sector. Panel C shows that the likelihood of being employed in the knowledge economy is similar among GCP participants and comparison students. We do find that GCP participants work less for tech manufacturing firms, perhaps due to their lower likelihood of graduating with engineering degrees. Thus, the results imply that GCP does not enhance gifted students' contribution to the knowledge economy.[23]

The discovery that GCP participation does not lead to improved career outcomes may initially come as a surprise. However, upon closer examination, this finding aligns with the evidence discussed throughout the paper. Firstly, we did not observe significant academic improvements in high school exit exam scores among GCP participants. Secondly, while we did find enhanced university outcomes, these may primarily reflect academic pursuits rather than financial gain motives, as evidenced by the pursuit of double majors and advanced degrees. Therefore, a plausible interpretation of these results is that the career choices made by GCP participants are not necessarily driven by a desire to maximize their financial returns based on their abilities.

5.5 Personal Outcomes

Finally, we also examine whether GCP participation affects personal outcomes. Panel A of Table 7 shows the results. We find no evidence for effects on marriage and fertility, but we do find a marginally significant increase in the likelihood of living outside Israel in 2020.[24] Panel B shows that GCP has a large positive effect on marrying a GCP participant, driven by matches within the class. It also increases the "quality" of the match, measured by the partner's UPET score. The effect on marriages with the same GCP participants is fascinating in light of the recent work by Kirkebøen et al. (2021), showing that colleges in Norway matter considerably for whom one marries by inducing

[22] In the main analysis, we focus on labor market outcomes in 2020, which is the latest year we observe. However, we also validate that the results are similar if we analyze labor-market outcomes during 2018–2020 (Appendix Table A4).

[23] We also examine changes in the likelihood of being employed in other common sectors among gifted students, such as the public, education, and health sectors. We do not find evidence for significant changes in any of these outcomes.

[24] We further validate that our estimates are not sensitive to different assumptions regarding the (missing) earnings of gifted students living abroad. For example, if we impute for these students the 10% (90%) percentile earnings of gifted students in our sample, the estimated effect of GCP on the earnings rank in 2020 is 0.003 (0.006) rank points with a standard error of 0.018 (0.018).

matches within educational institutions. Our findings suggest that GCPs also matter for the marital matches of gifted children.

This finding highlights a crucial aspect of the GCP's impact on gifted students' lives–the exposure to other high-achieving peers. To delve deeper into this matter, we extend our examination beyond marital considerations and explore professional connections. Specifically, we analyze outcomes at the plant level: the total number of workers and the total number of workers who are graduates of high school GCPs (excluding the individual itself).[25] The results, presented in Panel C of the table, affirm that GCP participants tend to work alongside peers who have also participated in similar programs, providing further evidence that the GCP significantly impacts the social connections of gifted students.

6 Validating the Findings' Causal Interpretation

We run many different exercises to examine directly different concerns regarding our methodology. In this section, we discuss them.

6.1 An Alternative and Complementary Analysis

In a supplementary analysis, we narrow our focus to the highest end of the ability distribution, using students' 8th-grade test scores. This approach increases the likelihood that students in the comparison group are indeed gifted. To construct the treatment group, we identify GCP participants who scored within the top 2 percentiles in any Meitzav 8th-grade tests. This treatment group comprises almost half of the GCP participants in our main sample, totaling 277 students.

For the comparison group, we encompass all students from other localities that lack access to a GCP, provided they also scored within the top 2 percentiles in any of the Meitzav 8th-grade tests. This results in a larger comparison group compared to our main analysis, comprising 4,786 students.

Utilizing these defined groups, we employ Equation 2 to estimate the differences in our primary outcomes between GCP participants and students from this expanded comparison group, conditional on pre-determined outcomes. The results are presented in column (2) of Table 8, and they closely resemble the findings from our main analysis (shown in column (1) of the same table).

[25] We exclude individuals working at plants with 500 workers or more, as this likely reflects large organizations rather than specific plants. However, the results remain robust even when including these individuals.

This supplementary analysis complements our main analysis by establishing a comparison group in a less prescriptive manner. By concentrating on the highest segments of the test score distributions, we can ensure that both the treatment and comparison group students are gifted. Furthermore, by controlling for their predetermined characteristics, we confirm that extraneous variables do not influence the results. Additionally, the students in this sample predominantly represent the upper end of the giftedness spectrum. As a result, the findings in this analysis reflect the impacts on the most gifted students.

The estimates in column (2) are strikingly similar. For example, the decline in degrees in engineering is 6.1pp in column (1) and 6.4pp in column (2). The increase in MA degrees is 9.1pp in column 1 and 10.4pp in column 2. The nill effect on the composite matriculation score is also identical, though insignificant. This similarity in the results is perhaps unsurprising, as we demonstrate that the effects of GCP are consistent across students with varying levels of giftedness. However, there is a big difference in the comparison group used in this alternative estimation, yet we find precisely the same results as in the main analysis. The sample size in column (2) is five times larger than that in column (1).

6.2 Placebo Analysis

We run a placebo analysis to support the causal interpretation of the results. Specifically, we define a new "treatment" group that includes non-gifted children who study in regular classrooms in other high schools in localities with a GCP. We randomly choose students from these localities to get treatment group size, which is similar to this employed in our main analysis. Then, we implement an identical matching algorithm to select a comparison group for these students and estimate the conditional difference in primary outcomes between both groups of students. Table 9 shows that this exercise yields desired null results concerning the primary outcomes, further supporting the validity of our analysis.[26]

6.3 Robustness for Specification Choices

Alternative matching algorithms. In the main estimation, we implement a matching algorithm and set a caliper of 0.1SD. This choice is somewhat arbitrary, but it does not affect the results. Columns (2)-(3) of Appendix Table A6 show that the results are similar when using other calipers (0.2SD and 0.05SD). Column (1) shows

[26] We also run the same exercise when using a "treatment" group of very talented students in localities with a GCP, who did not participate in the GCP. Appendix Table A5 shows the results, with small and mostly insignificant estimates.

the results of our main specification for comparison. We also validate that the results are identical if we match with replacement (column (4)). Finally, we also validate that the results are robust for changing the propensity score model. We use a gradient-boosting algorithm (Chen and Guestrin, 2016) instead of a logit model that allows a more flexible fit. We find similar results, as shown in column (5).

Choosing comparison students from alternative groups of localities. In the main analysis, we match students with equally gifted children from other localities with no GCP during our sample period. We also validate that the results are robust for choosing comparison group's students from only large cities (with an above-median number of students in the locality) and from localities with a GCP. The results are shown in columns (6)-(7) of Appendix Table A6.

6.4 Robustness for Changing the Ability and Motivation Measures

Excluding matriculation indicators from the matching specification. Our main specification includes matriculation indicators as measures for academic motivation. However, excluding these variables from the matching specification does not affect the main results. Column (2) of Table 10 shows that the main results are similar when excluding all indicators for elective matriculation subjects (including only indicators for five credits in math and English). The only difference from our main results is the small positive effects on the likelihood of achieving a degree, which is small in magnitude and may still be due to some students still studying toward their degree. The strong positive effect on the likelihood of a double major degree and the nill effects on labor-market outcomes remain identical.

Using 5^{th}-grade test scores as the ability measure. 5^{th}-grade Meitzav test scores are available for high school graduates of 2009 or later. We use the graduates between 2009 and 2013 to validate that our main results are robust for using these test scores as the ability measure in the matching specification. Column (3) of Table 10 shows that the results are similar to our main results. Note that the positive effect on the likelihood of achieving an undergraduate degree is not surprising since these students were younger in 2020 (25–29); thus, this estimate may reflect the tendency of GCP participants to complete their degrees earlier, as shown in our main results. Similarly, the marginally significant effect on employment in the knowledge economy may reflect that GCP participants tend to integrate into these sectors earlier.

Using UPET scores as the ability measure. An alternative ability measure that we can use is the UPET scores. Remarkably, while most students in Israel start their higher education studies at 22-23 years old, most GCP participants take their first university entrance test while still in high school and before the program's midpoint. This provides an opportunity to utilize these scores as an alternative measure of ability. Additionally, these tests are designed to assess intelligence and may not be influenced by participation in the GCP. To support this assertion, we present evidence that the UPET scores are not influenced by the age at which the test is taken. In Appendix Table A7, it can be observed that the differences in test scores between students who took the tests at earlier and later ages are relatively small. This holds particularly true for GCP participants and their quantitative scores (column (4) of the table), where all differences are statistically insignificant.

The finding that UPET scores remain consistent among highly talented individuals regardless of testing age may not be surprising, given that its structure and content closely resemble the SAT and CAT used in the U.S. for similar purposes. Moreover, research has demonstrated a high correlation between these tests and IQ and other measures of cognitive ability (Koenig et al., 2008; Beaujean et al., 2006; Frey and Detterman, 2004). Therefore, it is reasonable to expect minimal variation by age for the UPET, although the evidence for its correlation with IQ test scores is somewhat limited.

Importantly, this finding enables us to consider the UPET score as a predetermined variable, making it applicable to our matching specification. We demonstrate that the main results of our analysis remain consistent when utilizing UPET scores in place of the 8th-grade Meitzav test scores. We concentrate on the subset of early test-takers (those who took the UPET by age 17) and utilize only the quantitative UPET score as the measure of ability. As depicted in column (4) of Table 10, the results generally align with the main findings. These insights are crucial as they allow us to analyze the effects of the GCP for older cohorts, for whom the UPET scores represent the only available measure of ability in our dataset.

We also demonstrate the effectiveness of each ability measure in mitigating baseline disparities in the alternative measure. In Appendix Figure A1, we present the distributions of the UPET scores (in all domains) before and after the matching process, which uses the Meitzav test scores as the ability measure. Before matching, the UPET score distributions exhibit a notable advantage for GCP participants. However, post-matching, most of these differences are eliminated, with the disparities in the quantitative score distributions becoming only marginally statistically significant. Similarly, matching based on the quantitative UPET scores effectively mitigates baseline disparities in the 8th-grade Meitzav test scores, particularly in math, where the score distributions become statistically indistinguishable (Appendix Figure A2).

6.5 Persistence of the Results Until Advanced Ages

We extend our analysis to a sample of high school graduates from 1992 to 2005, providing insight into the long-term persistence of our findings (these individuals were ages 33-46 in 2020). For these students, the only ability measure available is their UPET scores. Focusing initially on those who took the UPET early and employing only quantitative UPET scores as the ability measure in our matching specification, column (1) of Table 11 confirms the persistent impact. We observe a significant effect on the likelihood of obtaining a double-major degree while finding no discernible effects on labor-market outcomes such as earnings rank or the probability of working in knowledge-producing sectors. Given the older age of these cohorts in 2020 compared to our main sample, we can also discern a substantial increase in the likelihood of attaining a Ph.D. degree (27%, significant at the 95% confidence level). We also show that the results are similar when extending the sample to all test-takers and using the scores in all three domains (Columns (2)-(4)).

Additionally, Figure 5 compares the earnings distributions of both groups of gifted students. Similar to what we found in the primary sample, there are no significant differences in earnings between GCP participants and the matched comparison students. If anything, the comparison group's students may earn a bit more, but this difference is statistically insignificant. The figure also shows the large gap between both groups of gifted students relative to non-gifted students. The overall impression is that gifted students do well in the labor market, regardless of their participation in a GCP.

7 Heterogeneity of the GCP Effects

In this section, we examine the heterogeneity of the effects. To estimate how the effects vary along dimension z, we estimate the following model:

$$Y_i = \alpha^h + X_i'\beta^h + \gamma^h \times z_i + \tau^h \times GCP_i + \delta^h \times GCP_i \times z_i + \varepsilon_i \tag{3}$$

The coefficient of interest is δ^h, which captures how the effects vary along z. We report the results regarding the heterogeneity of the effects on the main outcomes.

7.1 Gender

We begin by investigating potential gender differences in the impact of GCP participation. As shown in column (1) of Table 12, we present the estimated disparities in the effects of GCP on boys and girls for our main outcomes. We observe that only males drive the increase in the likelihood of obtaining a computer science degree or an MA degree. Nevertheless, most of the differences (δ^h) are statistically insignificant, suggesting that the GCP exerts a similar effect on both boys and girls. This holds for double majors, earnings, and marital match quality.

While these findings provide valuable insights, it's important to acknowledge a potential limitation–our sample size is relatively small, so we may only be able to detect substantial differences in the data. To address this concern, we extended our analysis to the sample of older cohorts, which is larger. The results in column (1) of Table 13 are generally consistent with those from our main sample. We did find a significant difference in the effect on earnings, indicating that females experienced smaller increases in earnings. However, it's important to note that the effects on earnings were statistically insignificant for both males and females, with estimated changes of 1.8 and -2.4 rank points, respectively, and standard errors of 1.3 and 1.8.

7.2 Socio-economic status (SES)

Another potentially important source of heterogeneity in GCP treatment effects is the participants' SES. To explore that, we estimate equation 3, defining z as an indicator for higher SES backgrounds, proxied by a father education of 15 years (the minimal number of years required to attain a degree) or more. Column (2) of Table 12 shows the estimated differences. The results show that the effects of GCP do not exhibit significant variation based on a student's SES background. Similarly, when we analyze the outcomes of the older cohorts, the differences remain insignificant (Column (2) of Table 13).[27] The insensitivity of the estimated effects of GCP to SES variation sharply contrasts with the effects of many other schooling inputs, which vary by student's background.

7.3 Giftedness

We also examine a model where we allowed for GCP impact heterogeneity by the giftedness level. We divided the sample into two groups based on their (pre-treatment) academic achievement, specifically by achieving a math test score of at least one standard deviation above the median in the 8^{th}-grade Meitzav test scores. The results are shown in column (3) of Table 12. We do not detect any significant difference in the effects. When analyzing the older cohorts and splitting the samples according to achievement in the quantitative UPET score, we find that the effects on a double major

[27] The only marginally significant estimate is the effect on earnings, but the effects on both groups are statistically insignificant.

degree are driven by students with lower levels of giftedness (Column (3) of Table 13).

7.4 Length of Participation in the GCP

A third of the GCP participants in the main sample were part of the program since middle school, while the remaining started the GCP in high school. This setup enables us to address a crucial aspect of GCP treatment effects heterogeneity: whether extended participation (from 7th to 12th grades) impacts student outcomes differently than shorter participation (from 10th to 12th grades). Our results in column (4) of Table 12 do not show significant heterogeneity. This suggests that GCP effects remain similar regardless of the duration of program participation. This analysis is limited to the main sample, as information on middle school GCP participation is unavailable for the older cohorts.

8 Conclusion

Gifted children receive special attention in many educational systems. With the growth of the knowledge economy, governments are becoming aware that nurturing gifted students is crucial for securing new generations of scientists, creators, and innovators. Yet, the vast majority of published research on the impact of GCP has only examined their effects on short-run outcomes, primarily by looking at their impact on standardized test scores and educational attainment. While important, a possibly more profound question of interest to society is the effect of such interventions on long-run life outcomes. We address this important question using Israel's unique setting, offering both wide-scope GCP and rich administrative data to follow program participants over their life cycle, from teenagerhood to adulthood, for some up to age 46.

We report several exciting and unique findings. First, no discernible effect of GCP on high school achievement. This finding is surprising given the abundance of educational inputs that GCP participants enjoy relative to the comparison group we use. We discuss two explanations for this finding. First, moving from an environment of 'big fish in a small pond' to being a 'big fish in a big pond' may cause anxiety and a decline in self-concept (which might translate into adverse effects on academic performance). Second, GCP's studies program incorporates advanced concepts and topics irrelevant to the standard curriculum. Thus, GCP participants may get educational benefits not manifested in higher matriculation test scores.

Secondly, we find large and significant effects on higher education outcomes. Among these, there is a large increase in the likelihood of graduating with a double major. This effect may reveal gifted children's multipotentiality and their difficulty selecting one area of interest to focus on. The higher likelihood of pursuing advanced degrees is consistent with the view that gifted children are pressured by parents and social circles to "maximize" their potential and not to "waste" it. As a result, we should not be surprised by our findings that there is no effect on earnings in adulthood, as the career path of gifted children is not necessarily guided by maximizing the financial return to their ability. Perhaps surprising is that GCP has no effect on integrating gifted children into work in sectors that produce knowledge. One explanation could be that gifted children are directed into these sectors in advance. Thus, the GCP plays no important role in these decisions.

Finally, we also find that participation in a GCP matters for marital matches and professional connections. These findings suggest that exposure to similarly gifted peers is an important benefit of GCP participation.

Against the benefits and gains accruing in gifted children's programs, we should note the potential loss to other students in the education system. Some evidence suggests that non-gifted children benefit from having high achievers and gifted children as peers (Lavy et al., 2012; Balestra et al., 2021). Thus, there is a concern that excluding gifted children from regular classes might negatively affect their peers.

References

Abadie, Alberto and Guido Imbens, "Simple and bias-corrected matching estimators for average treatment effects," 2002.

_ **and Guido W Imbens**, "On the failure of the bootstrap for matching estimators," *Econometrica*, 2008, *76* (6), 1537–1557.

_ **and Matias D Cattaneo**, "Econometric methods for program evaluation," *Annual Review of Economics*, 2018, *10*, 465–503.

Angrist, Joshua D, Victor Lavy, Jetson Leder-Luis, and Adi Shany, "Maimonides rule redux," *American Economic Review: Insights*, 2019, *1* (3), 309–324.

Balestra, Simone, Aurélien Sallin, and Stefan C Wolter, "High-ability influencers? The heterogeneous effects of gifted classmates," *Journal of Human Resources*, 2021, pp. 0920–11170R1.

Beaujean, A Alexander, Michael W Firmin, Andrew J Knoop, Jared D Michonski, Theodore P Berry, and Ruth E Lowrie, "Validation of the Frey and Detterman (2004) IQ prediction equations using the Reynolds Intellectual Assessment Scales," *Personality and Individual Differences*, 2006, *41* (2), 353–357.

Boettger, Eva Reid-Heiner and Eva Reid, "Gifted education in various countries of Europe," *Slavonic pedagogical studies journal*, 2015, *4* (2), 158–171.

Booij, Adam S, Ferry Haan, and Erik Plug, "Enriching students pays off: Evidence from an individualized gifted and talented program in secondary education," Technical Report 2016.

_ **,** _ **, and** _ , "Can gifted and talented education raise the academic achievement of all high-achieving students?," Technical Report 2017.

Bui, Sa A, Steven G Craig, and Scott A Imberman, "Is gifted education a bright idea? Assessing the impact of gifted and talented programs on students," *American Economic Journal: Economic Policy*, 2014, *6* (3), 30–62.

Card, David and Laura Giuliano, "Does gifted education work? For which students?," Technical Report, National Bureau of economic research 2014.

_ **and** _ , "Can tracking raise the test scores of high-ability minority students?," *American Economic Review*, 2016, *106* (10), 2783–2816.

Chen, Tianqi and Carlos Guestrin, "Xgboost: A scalable tree boosting system," *Proceedings of the 22nd ACM Sigkdd International Conference on Knowledge Discovery and Data Mining*, 2016, pp. 785–794.

Cohodes, Sarah R, "The long-run impacts of specialized programming for high-achieving students," *American Economic Journal: Economic Policy*, 2020, *12* (1), 127–66.

Dai, David Yun and Anne N Rinn, "The big-fish-little-pond effect: What do we know and where do we go from here?," *Educational Psychology Review*, 2008, *20*, 283–317.

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

Frey, Meredith C and Douglas K Detterman, "Scholastic assessment or g? The relationship between the scholastic assessment test and general cognitive ability," *Psychological science*, 2004, *15* (6), 373–378.

Herrmann, Julia, Isabelle Schmidt, Ursula Kessels, and Franzis Preckel, "Big fish in big ponds: Contrast and assimilation effects on math and verbal self-concepts of students in within-school gifted tracks," *British Journal of Educational Psychology*, 2016, *86* (2), 222–240.

Imbens, Guido W, "Nonparametric estimation of average treatment effects under exogeneity: A review," *Review of Economics and statistics*, 2004, *86* (1), 4–29.

Jr, James P Sampson and Ashley K Chason, "Helping gifted and talented adolescents and young adults: Make informed and careful career choices," in "Handbook of giftedness in children: Psychoeducational theory, research, and best practices," Springer, 2008, pp. 327–346.

Kerr, Barbara A and Nicholas Colangelo, "The college plans of academically talented students," *Journal of Counseling & Development*, 1988, *67* (1), 42–48.

Kerr, Barbara and Cheryl Erb, "Career counseling with academically talented students: Effects of a value-based intervention.," *Journal of Counseling Psychology*, 1991, *38* (3), 309.

Kirkebøen, Lars, Edwin Leuven, and Magne Mogstad, "College as a marriage market," Technical Report, National Bureau of Economic Research 2021.

Koenig, Katherine A, Meredith C Frey, and Douglas K Detterman, "ACT and general cognitive ability," *Intelligence*, 2008, *36* (2), 153–160.

Lavy, Victor, "Expanding school resources and increasing time on task: Effects on students' academic and noncognitive outcomes," *Journal of the European Economic Association*, 2020, *18* (1), 232–265.

_ **and Analia Schlosser**, "Mechanisms and impacts of gender peer effects at school," *American Economic Journal: Applied Economics*, 2011, *3* (2), 1–33.

_ **, Olmo Silva, and Felix Weinhardt**, "The good, the bad, and the average: Evidence on ability peer effects in schools," *Journal of Labor Economics*, 2012, *30* (2), 367–414.

Leung, S Alvin, Collie W Conoley, and Michael J Scheel, "The career and educational aspirations of gifted high school students: A retrospective study," *Journal of Counseling & Development*, 1994, *72* (3), 298–303.

Marsh, Herbert W and John W Parker, "Determinants of student self-concept: Is it better to be a relatively large fish in a small pond even if you don't learn to swim as well?," *Journal of personality and social psychology*, 1984, *47* (1), 213.

— **and Rhonda G Craven**, "The Pivotal Role of Frames of Reference in Academic Self-Concept Formation: The" Big Fish-Little Pond" Effect.," 2002.

— **, Marjorie Seaton, Ulrich Trautwein, Oliver Lüdtke, Kit-Tai Hau, Alison J O'Mara, and Rhonda G Craven**, "The big-fish–little-pond-effect stands up to critical scrutiny: Implications for theory, methodology, and future research," *Educational psychology review*, 2008, *20*, 319–350.

Mönks, Franz J, Robin Pflüger, and Radboud Universiteit Nijmegen, *Gifted education in 21 European countries: Inventory and perspective*, Citeseer, 2005.

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

Pfeiffer, SI, P Olszewski-Kubilius, L Limburg-Weber, and SI Pfeiffer, "Psychological considerations in raising a healthy gifted child," *Early gifts: Recognizing and nurturing children's talents*, 2003, pp. 173–185.

Plucker, Jonathan A, Nancy M Robinson, Thomas S Greenspon, John F Feldhusen, D Betsy McCoach, and Rena F Subotnik, "It's not how the pond makes you feel, but rather how high you can jump.," 2004.

Preckel, Franzis, Moshe Zeidner, Thomas Goetz, and Esther Jane Schleyer, "Female 'big fish'swimming against the tide: The 'big-fish-little-pond effect'and gender-ratio in special gifted classes," *Contemporary Educational Psychology*, 2008, *33* (1), 78–96.

Redding, Christopher and Jason A Grissom, "Do students in gifted programs perform better? Linking gifted program participation to achievement and nonachievement outcomes," *Educational Evaluation and Policy Analysis*, 2021, *43* (3), 520–544.

Rinn, Anne N, "Effects of programmatic selectivity on the academic achievement, academic self-concepts, and aspirations of gifted college students," *Gifted Child Quarterly*, 2007, *51* (3), 232–245.

Robins, James M and Ya'acov Ritov, "Toward a curse of dimensionality appropriate (CODA) asymptotic theory for semi-parametric models," *Statistics in medicine*, 1997, *16* (3), 285–319.

Robinson, Nancy M, Robin Gaines Lanzi, Richard A Weinberg, Sharon Landesman Ramey, and Craig T Ramey, "Family factors associated with high academic competence in former Head Start children at third grade," *Gifted Child Quarterly*, 2002, *46* (4), 278–290.

Rosenbaum, Paul R and Donald B Rubin, "The central role of the propensity score in observational studies for causal effects," *Biometrika*, 1983, *70* (1), 41–55.

Sajjadi, Seyed Hossein, F Gillian Rejskind, and Bruce M Shore, "Is multipotentiality a problem or not? A new look at the data," *High Ability Studies*, 2001, *12* (1), 27–43.

Tymms, Peter, "A test of the big fish in a little pond hypothesis: An investigation into the feelings of seven-year-old pupils in school," *School Effectiveness and School Improvement*, 2001, *12* (2), 161–181.

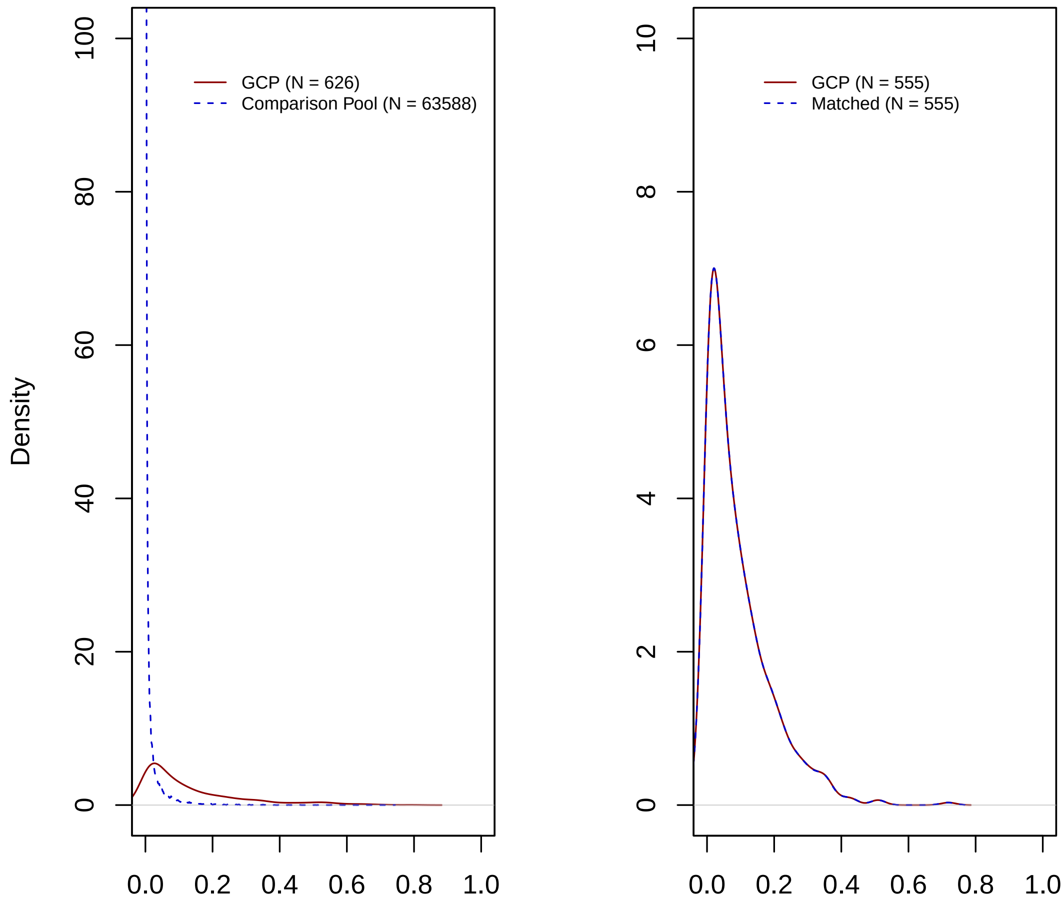
Vrignaud, Pierre, Denis Bonora, and Annie Dreux, "Counselling the gifted and talented in France: minimizing gift and maximizing talent," *International Journal for the advancement of counselling*, 2005, *27*, 211–228.

Zeidner, Moshe and Esther Jane Schleyer, "The big-fish–little-pond effect for academic self-concept, test anxiety, and school grades in gifted children," *Contemporary educational psychology*, 1999, *24* (4), 305–329.

_ and Inbal Shani-Zinovich, "A comparison of multiple facets of self-concept in gifted vs. non-identified Israeli students," *High Ability Studies*, 2015, *26* (2), 211–226.

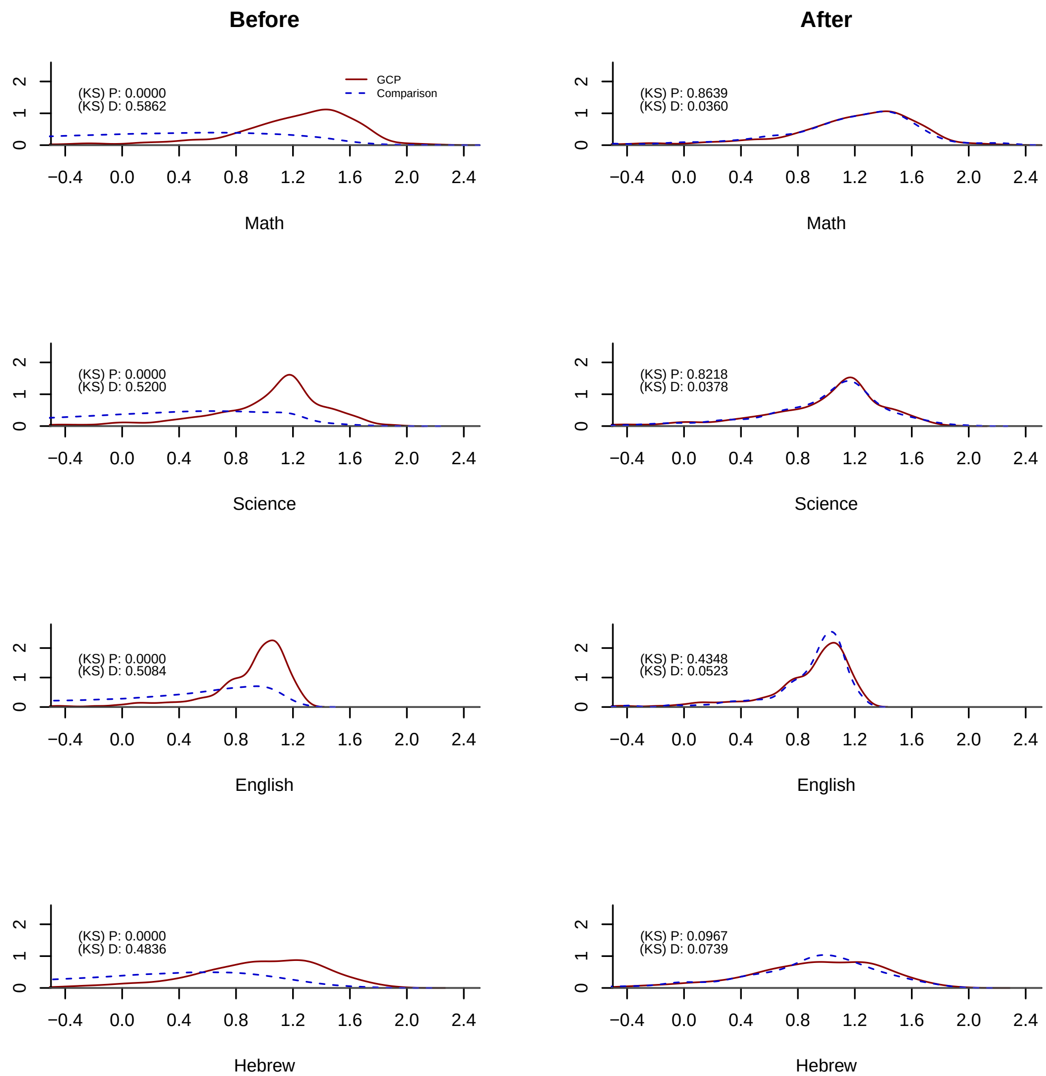
_ , _ , Gerald Matthews, and Richard D Roberts, "Assessing emotional intelligence in gifted and non-gifted high school students: Outcomes depend on the measure," *Intelligence*, 2005, *33* (4), 369–391.

Figure 1: Estimated Propensity Score Distributions for GCP Participation, Before and After Matching



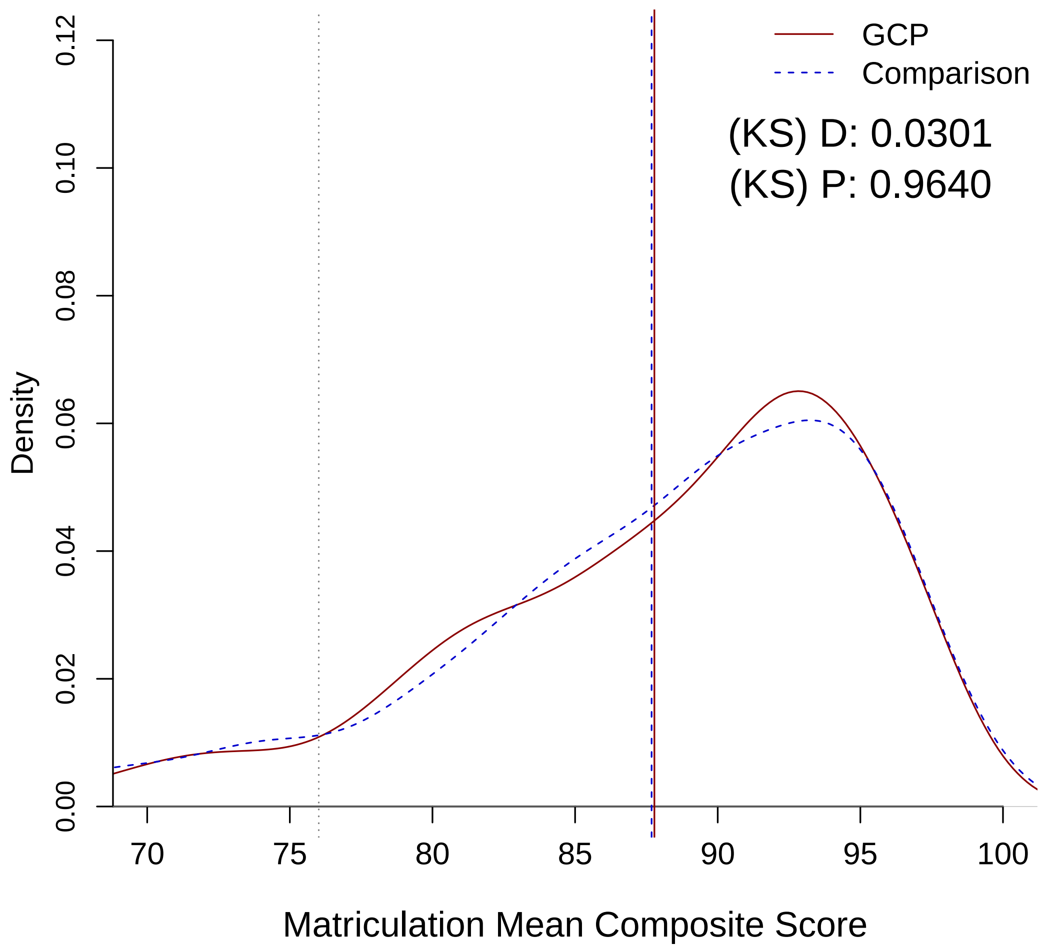
Notes: This figure plots the propensity score distributions for GCP participation. The solid red line represents the sample of GCP students, while the blue dashed line represents the comparison group, which includes non-GCP students from other cities. The left panel shows the distributions before matching, and the right panel shows the distributions after matching. The sample includes students who participated in the Meitzav middle school test during their 8th grade, constituting about half of the students in cohorts of high-school graduates from 2006 to 2010. GCP stands for Gifted Children's Program.

Figure 2: Pre-Treatment Middle-School Test Scores, Before and After Matching



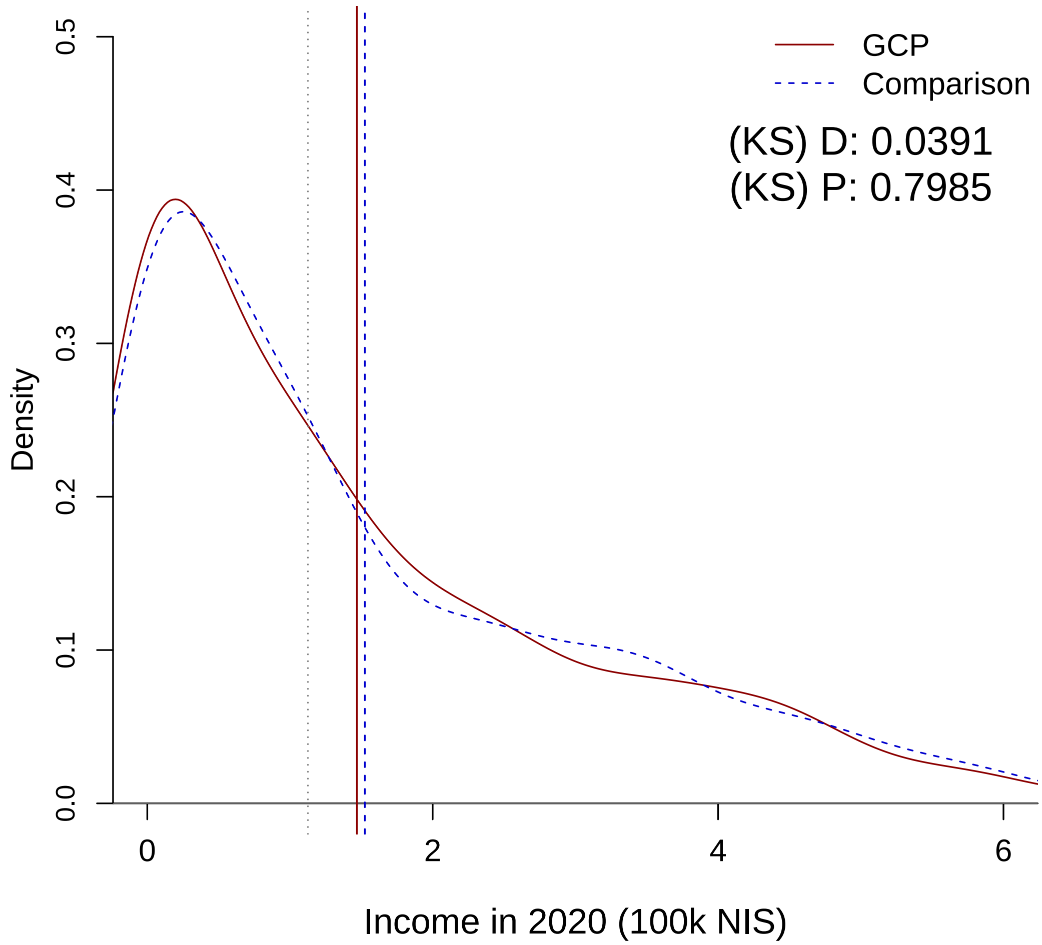
Notes: This figure plots the pre-treatment 8th-grade Meitzav test score distributions. The solid red lines represent the sample of GCP students, and the blue dashed lines represent the comparison group (which includes non-GCP students from other cities). The figures on the left show the distributions before the matching, and those on the right show the distributions after the matching. The figures also show the Kolmogorov–Smirnov test for the equality of the probability distributions. The sample includes students who participated in the 8th-grade Meitzav tests, about half of the students in cohorts of high-school graduates in 2006-2010. GCP stands for Gifted Children's Program.

Figure 3: Mean Composite Matriculation Score, GCP Participants and Comparison Group



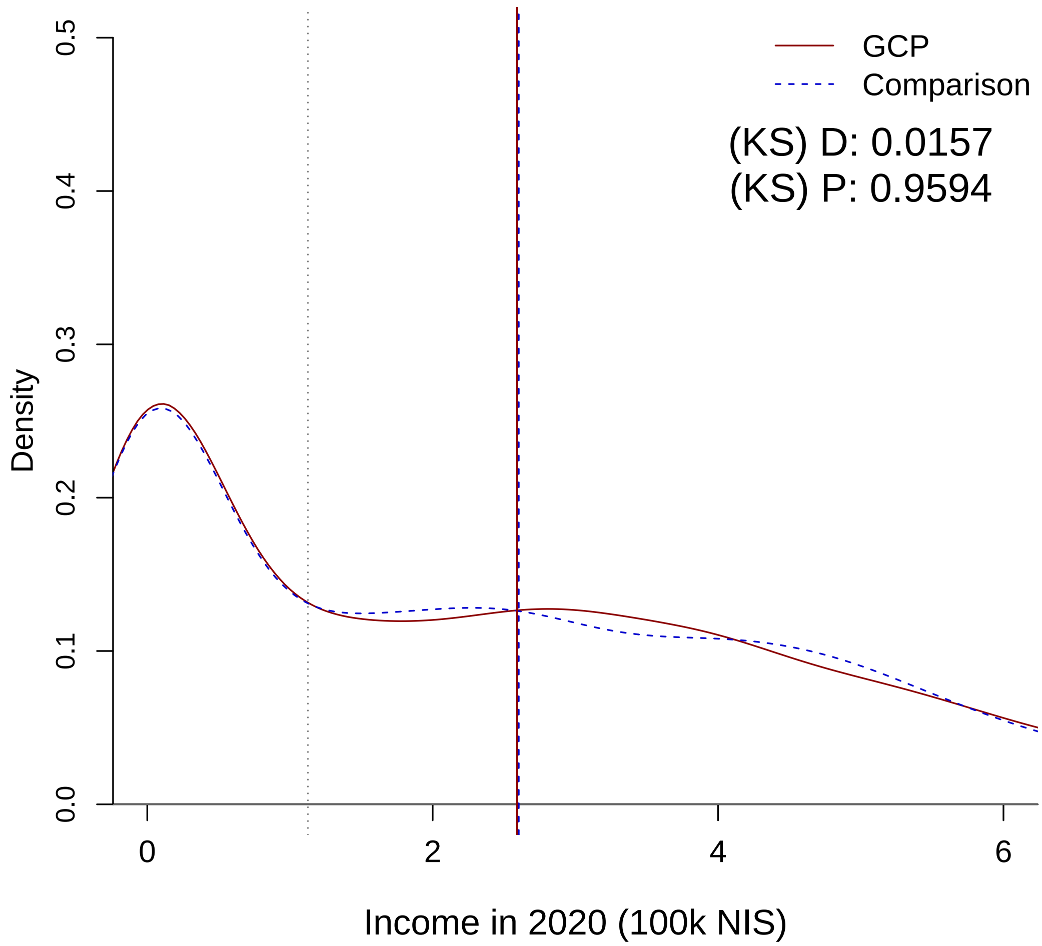
Notes: This figure plots the mean composite matriculation score distributions of GCP participants and the matched comparison group. The solid red line represents the sample of GCP students, and the blue dashed line represents the matched comparison group. The figure reports the Kolmogorov–Smirnov test for the equality of the probability distributions. The vertical lines represent the averages. The dotted grey line represents the average among students in the comparison group's pool, including all students from cities with no GCPs. The figure is based on the sample of students who participated in the 8th-grade Meitzav tests, about half of the students in cohorts of high-school graduates in 2006-2010. GCP stands for Gifted Children's Program.

Figure 4: Annual Earnings, GCP Participants and Comparison Group



Notes: This figure plots the distribution of the annual earnings in 2020 of GCP participants and the matched comparison group. The solid red line represents the sample of GCP students, and the blue dashed line represents the matched comparison group. The figure reports the Kolmogorov–Smirnov test for the equality of the probability distributions. The vertical lines represent the averages. The dotted grey line represents the average among students in the comparison group's pool, including all students from cities with no GCPs. The figure is based on the sample of students who participated in the 8th-grade Meitzav tests, about half of the students in cohorts of high-school graduates in 2006-2010. GCP stands for Gifted Children's Program.

Figure 5: Annual Earnings, GCP Participants and Comparison Group, 1992-2005 Graduates



Notes: This figure plots the distribution of the annual earnings in 2020 of GCP participants and the matched comparison group. The solid red line represents the sample of GCP students, and the blue dashed line represents the comparison group (which includes non-GCP students from other cities). The figure also shows the Kolmogorov–Smirnov test for the equality of the probability distributions. The dotted black line represents the comparison group pool, which includes non-gifted students. The sample includes students who participated in the UPET (until the age of 17) from the cohorts of high-school graduates in 1992-2005.

Table 1: Balancing Table, Parental Characteristics

	Comparison	GCP	Difference
	(1)	(2)	(3)
A. Annual Earnings (100K NIS)			
Father	1.65	1.53	-0.11
			(0.11)
Mother	0.78	0.85	0.07
			(0.06)
B. Years of Education			
Father	15.07	15.01	-0.06
			(0.17)
Mother	15.17	14.94	-0.23
			(0.17)
C. % Born in Israel			
Father	63.60	59.28	-4.32
			(2.92)
Mother	65.95	64.14	-1.80
			(2.86)
D. Age at Birth			
Father	32.19	32.62	0.43
			(0.31)
Mother	29.15	29.56	0.41
			(0.30)
Students	555	555	1,110
Schools	144	11	155

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

Notes: The table shows the balance between GCP participants and the matched comparison group in parental characteristics. The sample includes only students who participated in the 8th-grade Meitzav tests, about half of the students in cohorts of high-school graduates in 2006-2010. Annual earnings (panel A) refer to the total earnings earned in 2003, measured in 100K NIS. Standard errors are clustered at the school level. Columns (1) and (2) show the means among the matched comparison group and the treatment group (GCP participants). Column (3) shows the outcome's unconditional difference and its standard error. GCP stands for Gifted Children's Program.

Table 2: Balancing Table, Student Characteristics

	Comparison	GCP	Difference
	(1)	(2)	(3)
A. Background			
Number of Siblings	1.72	1.68	-0.04
			(0.08)
Family Order	1.71	1.78	0.07
			(0.06)
% Born in Israel	85.05	82.52	-2.52
			(2.21)
B. Matriculation Program			
Total Credits	29.23	29.74	0.50
			(0.36)
% Math, 5 Credits	75.32	75.32	0.00
			(2.59)
% English, 5 Credits	90.27	91.35	1.08
			(1.74)
% Physics, 5 Credits	48.65	48.29	-0.36
			(3.00)
% Computer Science,	48.47	49.01	0.54
			(3.00)
% Chemistry, 5 Credits	30.09	28.65	-1.44
			(2.74)
% Biology, 5 Credits	15.14	15.50	0.36
			(2.16)
Students	555	555	1,110
Schools	144	11	155

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

Notes: The table shows the balance between GCP participants and the matched comparison group in their personal characteristics. The sample includes only students who participated in 8th-grade Meitzav tests, about half of the students in cohorts of high-school graduates in 2006-2010. Standard errors are clustered at the school level. Columns (1) and (2) show the means among the matched comparison group and the treatment group (GCP participants). Column (3) shows the outcome's unconditional difference and its standard error. GCP stands for Gifted Children's Program.

Table 3: The Impact of GCP Participation on the Class Environment

	Comparison	GCP	Estimated Effect
	(1)	(2)	(3)
Number of Students	34.21	27.47	-6.74***
			(0.29)
% Males	52.69	61.94	9.24***
			(0.71)
Father Education	14.21	15.56	1.35***
			(0.06)
Mother Education	14.32	15.35	1.02***
			(0.06)
% 5 Credits in Physics	24.05	50.53	26.47***
			(1.07)
UPET Score	586.06	658.58	72.52***
			(2.43)
% BA	64.39	78.23	13.83***
			(0.70)
% MA	13.28	26.36	13.08***
			(0.48)
Students	555	555	1,110
Schools	144	11	155

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

Notes: The table shows estimates for the impact of GCP participation on high school class-level outcomes. The sample includes only students who participated in 8th-grade Meitzav tests, about half of the students in cohorts of high-school graduates in 2006-2010. Columns (1) and (2) show the means in class-level outcomes among the matched comparison group and the treatment group (GCP participants). Column (3) shows the conditional difference (τ from equation (2)) in the outcome and its standard error (clustered at the school level). GCP stands for Gifted Children's Program.

Table 4: The Impact of GCP on Higher Education Outcomes

	Comparison	GCP	Estimated Effect
	(1)	(2)	(3)
A. BA Degrees			
% Enrollment, Until Age 18	2.52	27.42	24.90***
			(1.98)
% Enrollment	90.81	92.64	1.83
			(1.49)
% Attainment	76.40	79.68	3.29
			(2.33)
B. MA Degrees			
% Enrollment	26.67	36.88	10.22***
			(2.78)
% Attainment	18.92	27.96	9.05***
			(2.55)
C. PHD Degrees			
% Enrollment	3.60	6.72	3.12**
			(1.30)
% Attainment	0.18	1.21	1.03*
			(0.54)
D. MED Degrees			
% Enrollment	4.86	4.99	0.13
			(1.33)
% Attainment	1.80	3.04	1.24
			(0.99)
Students	555	555	1,110
Schools	144	11	155

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

Notes: The table shows estimates for the impact of GCP participation on higher education outcomes. The sample includes only students who participated in 8th-grade Meitzav tests, about half of the students in cohorts of high-school graduates in 2006-2010. Columns (1) and (2) show the means among the matched comparison group and the treatment group (GCP participants). Column (3) shows the conditional difference (τ from equation (2)) in the outcome and its standard error (clustered at the school level). GCP stands for Gifted Children's Program.

Table 5: The Impact of GCP on Undergraduate Degree Fields of Study

	Comparison	GCP	Estimated Effect
	(1)	(2)	(3)
A. STEM Fields			
% Any STEM	46.31	44.90	-1.41
			(2.74)
% Math, Computer Science, Statistics	15.86	20.66	4.80**
			(2.22)
% Engineering	23.24	17.19	-6.05**
			(2.38)
% Physical Sciences	7.75	9.10	1.35
			(1.65)
% Biological Sciences	4.68	4.39	-0.30
			(1.23)
B. Double Majors			
% Any Double Major	18.74	24.58	5.84**
			(2.51)
% Double STEM	6.85	8.88	2.04
			(1.61)
% STEM & Other	3.24	3.10	-0.15
			(1.07)
% Double Non-STEM	8.83	12.54	3.71**
			(1.85)
Students	555	555	1,110
Schools	144	11	155

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

Notes: The table shows estimates for the impact of GCP participation on fields of study in undergraduate degrees. The sample includes only students who participated in 8th-grade Meitzav tests, about half of the students in cohorts of high-school graduates in 2006-2010. Columns (1) and (2) show the means among the matched comparison group and the treatment group (GCP participants). Column (3) shows the conditional difference (τ from equation (2)) in the outcome and its standard error (clustered at the school level). GCP stands for Gifted Children's Program.

Table 6: The Impact of GCP on Labor Market Outcomes

	Comparison	GCP	Estimated Effect
	(1)	(2)	(3)
A. Annual Earnings			
Earnings (100K NIS)	1.52	1.45	-0.08
			(0.09)
Log Earnings	11.62	11.66	0.04
			(0.09)
% Earnings Rank	58.10	57.55	-0.55
			(1.76)
% Top 10% Earners	23.96	22.89	-1.07
			(2.47)
B. Employment			
% Salaried Employment	81.44	78.91	-2.53
			(2.44)
% Self Employment	6.13	4.63	-1.50
			(1.38)
C. Employment in Knowledge Producing Sectors			
% Any Knowledge Sector	41.26	41.19	-0.07
			(2.86)
% Tech Services	30.45	32.33	1.88
			(2.70)
% Tech Manufacturing	5.77	1.88	-3.89***
			(1.18)
% Academic	5.05	6.98	1.93
			(1.46)
Students	555	555	1,110
Schools	144	11	155

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

Notes: The table shows estimates for the impact of GCP participation on labor-market outcomes in 2020. The sample includes only students who participated in 8th-grade Meitzav tests, about half of the students in cohorts of high-school graduates in 2006-2010. Columns (1) and (2) show the means among the matched comparison group and the treatment group (GCP participants). Column (3) shows the conditional difference (τ from equation (2)) in the outcome and its standard error (clustered at the school level). GCP stands for Gifted Children's Program.

Table 7: The Impact of GCP on Personal, Partner, and Coworkers Outcomes

	Comparison	GCP	Estimated Effect
	(1)	(2)	(3)
A. Personal Outcomes			
% Outside Israel	4.14	6.65	2.51*
			(1.39)
% Marriage	34.77	37.31	2.54
			(2.78)
% Has Kids	17.30	18.72	1.42
			(2.13)
B. Partner Outcomes			
% GCP Participant	1.66	8.84	7.18***
			(2.48)
UPET Total Score	581.68	603.78	22.10*
			(12.04)
Annual Earnings (100K NIS)	1.35	1.28	-0.07
			(0.15)
C. Coworkers Outcomes			
Total	125.83	136.93	11.10
			(14.34)
GCP Participants	1.46	2.49	1.03***
			(0.36)
Students	555	555	1,110
Schools	144	11	155

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

Notes: The table shows estimates for the impact of GCP participation on personal, partner, and coworker outcomes. The sample includes only students who participated in 8th-grade Meitzav tests, about half of the students in cohorts of high-school graduates in 2006-2010. Columns (1) and (2) show the means among the matched comparison group and the treatment group (GCP participants). Column (3) shows the conditional difference (τ from equation (2)) in the outcome and its standard error (clustered at the school level). GCP stands for Gifted Children's Program, and UPET stands for University Psychometric Entrance Test.

Table 8: The Impact of GCP on Outcomes, Alternative Identification

	Main	Alternative
	(1)	(2)
A. Matriculation		
Mean Composite Score	-0.19	-0.22
	(0.33)	(0.47)
B. Higher Education		
% BA Attainment	3.29	2.94
	(2.33)	(2.14)
% BA, Double Major	5.84**	9.15***
	(2.51)	(2.93)
% BA, Math, CS, Stats	4.80**	5.73*
	(2.22)	(3.29)
% BA, Engineering	-6.05**	-6.38***
	(2.38)	(2.27)
% MA Attainment	9.05***	10.24**
	(2.55)	(3.66)
C. Labor-market		
% Earnings Rank	-0.55	-2.35
	(1.76)	(2.32)
% Knowledge Economy	-0.07	-2.49
	(2.86)	(2.74)
D. Personal		
% Married	2.54	-0.65
	(2.78)	(2.74)
Partner UPET Score	22.10*	34.99***
	(12.04)	(8.14)
Students	1,110	5,063

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

Notes: The table compares our estimates for the impact of GCP participation using the main methodology described throughout the paper (column (1)) and the complementary approach described in this appendix section (column (2)). In the complementary analysis, the comparison group includes students from the top score distribution in localities with no access to a GCP. The baseline sample includes only students who participated in 8th-grade Meitzav tests, about half of the students in cohorts of high-school graduates in 2006-2010. The numbers represent the conditional difference (τ from equation (2)) in the outcome and its standard error (clustered at the school level).

Table 9: Placebo Exercise, The "Impact" of Regular Classes on Outcomes

	Comparison	GCP	Estimated Effect
	(1)	(2)	(3)
A. Matriculation			
Mean Composite Score	76.09	75.80	-0.29
			(0.51)
B. Higher Education			
% BA Attainment	48.80	51.29	2.48
			(2.56)
% BA, Double Major	8.29	7.23	-1.06
			(1.60)
% BA, Math, CS, Stats	3.68	4.75	1.07
			(1.15)
% BA, Engineering	9.21	8.91	-0.29
			(1.59)
% MA Attainment	9.21	10.07	0.86
			(1.66)
C. Labor Market			
% Earnings Rank	51.32	53.42	2.10
			(1.37)
% Knowledge-Economy	16.76	15.89	-0.86
			(2.13)
D. Personal			
% Married	48.43	45.06	-3.37
			(2.93)
Partner UPET Score	560.79	557.74	-3.05
			(13.21)
Students	543	543	1,086
Schools	194	110	304

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

Notes: The table shows the results of a placebo exercise, estimating the effects of studying in regular classes (in high schools with no GCP, located in localities with a GCP), using the same matching algorithm we use throughout our analysis. The sample includes only students who participated in 8th-grade Meitzav tests, about half of the students in cohorts of high-school graduates in 2006-2010. Columns (1) and (2) show the means among the matched comparison group (students in other localities) and the treatment group (randomly selected students from regular classes in localities with a GCP). Column (3) shows the conditional difference in the outcome (τ from equation (2)) and its standard error (clustered at the school level). We find null results, as expected, supporting the validity of our design. GCP stands for Gifted Children's Program.

Table 10: Robustness of the Results, Alternative Proxies for Ability and Motivation

	(1)	(2)	(3)	(4)
A. Matriculation				
Mean Composite Score	-0.19	0.23	0.44	0.12
	(0.33)	(0.31)	(0.31)	(0.28)
B. Higher Education				
% BA Attainment	3.29	4.56**	9.03***	0.77
	(2.33)	(2.25)	(2.63)	(2.09)
% BA, Double Major	5.84**	7.31***	5.24**	5.84**
	(2.51)	(2.35)	(1.93)	(2.55)
% BA, Math, CS, Stats	4.80**	10.42***	9.37***	4.13*
	(2.22)	(2.04)	(2.06)	(2.29)
% BA, Engineering	-6.05**	-8.16***	-4.71**	-8.05***
	(2.38)	(2.31)	(1.91)	(2.61)
% MA Attainment	9.05***	10.31***	5.32***	3.49
	(2.55)	(2.37)	(1.79)	(2.71)
C. Labor-market				
% Earnings Rank	-0.55	-1.75	2.16	-3.43*
	(1.76)	(1.67)	(1.55)	(1.84)
% Knowledge-Economy	-0.07	2.24	5.20*	-2.66
	(2.86)	(2.72)	(2.67)	(2.94)
D. Personal				
% Married	2.54	2.12	-1.34	-1.49
	(2.78)	(2.69)	(2.04)	(2.86)
Partner UPET Score	22.10*	16.15*	-3.16	18.80*
	(12.04)	(9.56)	(27.74)	(9.55)
Cohorts	06-10	06-10	09-13	06-10
Ability Measure	8th-grade	8th-grade	5th-grade	UPET
Matriculation Program	+		+	+
Number of observations	1,110	1,222	1,192	1,142

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

Notes: The table shows estimates for the impact of GCP participation on outcomes, using different sets of variables in the matching specification. Column (1) shows our main sample and specification. Column (2) shows the results excluding the indicators for elective Matriculation subjects (including only the mandatory math and English subjects). Column (3) shows the results using 5th-grade Meitzav test scores as the ability measure. Column (4) shows the results using quantitative UPET scores (among those tested by the age of 17). The sample includes only students who participated in the relevant tests from the cohorts mentioned in the table. Each column shows the conditional difference (τ from equation (2)) in the outcome mentioned on the left and its standard error (clustered at the school level). GCP stands for Gifted Children's Program, and UPET stands for University Psychometric Entrance Test.

Table 11: Robustness of the Results, Sample of 1992–2005 Graduates

	(1)	(2)	(3)	(4)
A. Higher Education				
% BA Attainment	-0.44	-0.50	-0.44	-1.37*
	(0.85)	(0.85)	(0.85)	(0.69)
	91.91	92.08	91.91	90.93
% BA, Double Major	3.13**	4.33***	3.13**	2.32**
	(1.40)	(1.40)	(1.40)	(1.05)
	28.03	26.65	28.03	27.52
% BA, Math, CS, Stats	1.15	1.70	1.15	1.70*
	(1.23)	(1.23)	(1.23)	(0.87)
	21.66	21.20	21.66	17.90
% BA, Engineering	-6.80***	-5.40***	-6.80***	-5.22***
	(1.31)	(1.31)	(1.31)	(0.95)
	29.65	28.20	29.65	25.51
% MA Attainment	2.58*	1.36	2.58*	2.85**
	(1.54)	(1.55)	(1.54)	(1.15)
	46.45	46.98	46.45	41.86
% Ph.D. Attainment	2.47**	2.25**	2.47**	2.87***
	(0.94)	(0.94)	(0.94)	(0.65)
	9.09	9.17	9.09	7.14
B. Labor-market				
% Earnings Rank	0.34	-0.69	0.34	-1.91**
	(1.03)	(1.03)	(1.03)	(0.75)
	67.62	68.28	67.62	67.98
% Knowledge-Economy	-1.19	-1.23	-1.19	-0.29
	(1.49)	(1.50)	(1.49)	(1.10)
	40.65	40.95	40.65	36.79
C. Personal				
% Married	-1.63	-1.11	-1.63	-2.74**
	(1.43)	(1.44)	(1.43)	(1.09)
	68.47	68.84	68.47	68.35
Partner UPET Score	14.59***	11.26***	14.59***	11.01***
	(3.63)	(3.61)	(3.63)	(2.79)
	604.69	607.23	604.69	606.38
Only quantitative score	+		+	
Only early takers	+	+		
Number of observations	4,202	4,142	4,202	7,252

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

Notes: The table shows estimates for the impact of GCP participation on outcomes using different specifications and samples. The sample includes only students who participated in the UPET from the cohorts of high-school graduates between 1992–2005. In columns (1) and (3) the matching specification includes only the quantitative UPET scores as the ability measure, and in columns (2) and (4), it includes UPET scores in all three domains. In columns (1) and (2) the sample is restricted to students who took the UPET early (until the age of 17), and in columns (3) and (4), it includes students who took the UPET at any age. Each column shows the conditional difference (τ from equation (2)) in the outcome mentioned on the left and its standard error (clustered at the school level). Below them, we also report the baseline mean in the outcome (i.e., the average among the comparison group's students). GCP stands for Gifted Children's Program, and UPET stands for University Psychometric Entrance Test.

Table 12: Heterogeneity of The Impact of GCP

	Female	High SES	High ability	Since midschool
	(1)	(2)	(3)	(4)
A. Matriculation				
Mean Composite Score	0.15	0.71	0.17	-0.67
	(0.64)	(0.70)	(0.65)	(0.48)
B. Higher Education				
% BA Attainment	-3.52	3.42	-3.51	-3.14
	(4.70)	(4.97)	(4.79)	(3.48)
% BA, Double Major	-4.63	-1.46	4.90	5.08
	(5.20)	(5.16)	(4.75)	(4.19)
% BA, Math, CS, Stats	-13.83***	0.19	0.68	3.75
	(4.18)	(4.47)	(3.86)	(3.51)
% BA, Engineering	5.35	3.07	-6.70	-7.36**
	(4.64)	(4.87)	(4.28)	(3.43)
% MA Attainment	-11.57**	3.48	-2.89	1.67
	(5.28)	(5.23)	(5.10)	(4.30)
C. Labor Market				
% Earnings Rank	2.48	-0.18	1.05	3.33
	(3.51)	(3.61)	(3.32)	(2.87)
% Knowledge-Economy	-6.18	5.39	-1.04	5.91
	(5.76)	(5.91)	(5.44)	(4.61)
D. Personal				
% Married	8.05	2.18	4.42	0.47
	(5.89)	(5.78)	(5.27)	(4.31)
Partner UPET Score	-5.40	13.15	10.62	26.94
	(22.85)	(19.23)	(21.61)	(16.49)

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

Notes: The table shows estimates for the heterogeneity of the impact of GCP participation on outcomes. The sample includes only students who participated in 8th-grade Meitzav tests, about half of the students in cohorts of high-school graduates in 2006-2010. Columns (1)-(4) show the estimated difference between the effects on different groups of participants according to the characteristics mentioned at the top row (δ^h from equation (3)) and its standard error (clustered at the school level). GCP stands for Gifted Children's Program, and UPET stands for University Psychometric Entrance Test.

Table 13: Heterogeneity of The Impact of GCP, Sample of 1992–2005 Graduates

	Female	High SES	High ability
	(1)	(2)	(3)
A. Higher Education			
% BA Attainment	1.01	1.04	0.97
	(1.69)	(1.68)	(1.67)
% BA, Double Major	4.54	-0.37	-5.66**
	(2.99)	(2.83)	(2.64)
% BA, Math, CS, Stats	-2.82	-0.52	0.54
	(2.37)	(2.47)	(2.06)
% BA, Engineering	2.22	0.36	2.34
	(2.61)	(2.61)	(2.26)
% MA Attainment	-5.42*	3.67	4.02
	(3.25)	(3.10)	(2.87)
% Ph.D. Attainment	-0.74	1.06	1.38
	(1.97)	(1.88)	(1.65)
B. Labor Market			
% Earnings Rank	-4.14**	-3.46*	2.05
	(2.09)	(2.06)	(1.87)
% Knowledge-Economy	-3.72	-4.09	0.27
	(3.06)	(2.99)	(2.71)
C. Personal			
% Married	-5.88**	2.33	4.24
	(2.99)	(2.88)	(2.69)
Partner UPET Score	2.65	2.58	-1.08
	(7.66)	(6.66)	(6.95)

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

Notes: The table shows estimates for the heterogeneity of the impact of GCP participation on outcomes. The sample includes only students who participated in the UPET early (until the age of 17) from the cohorts of high-school graduates between 1992 and 2005. Columns (1)-(3) show the estimated difference between the effects on different groups of participants according to the characteristics mentioned at the top row (δ^h from equation (3)) and its standard error (clustered at the school level).

Online Appendices

Appendix A Appendix

A.1 Data Sources

We use several panel datasets from Israel's Central Bureau of Statics (CBS). CBS allows restricted access to this data in their protected research lab. The underlying data sources include the following. The population registry data consists of a fictitious individual national I.D. 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 marital status, number of children, and birth year. In addition, administrative records of the Ministry of Education on Israeli high schools' universe during the 1992-2016 school years provide the following student's family-background variables: parental schooling, number of siblings, country of birth, ethnicity, student's detailed 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. Another source is the Higher Council of Education records of post-secondary completed degrees (B.A., MA, and Ph.D.), the institution of study (colleges and universities), majors (one or two), and completion date. Finally, we also observe Israel Tax Authority (ITA) information on income and earnings of employees and self-employed individuals for 2000–2020, and three-digit code of industry of employment. CBS matched and merged these files using the individual-level national I.D. number. The matching is perfect without the loss of observations.

A.2 Identifying GCP classes

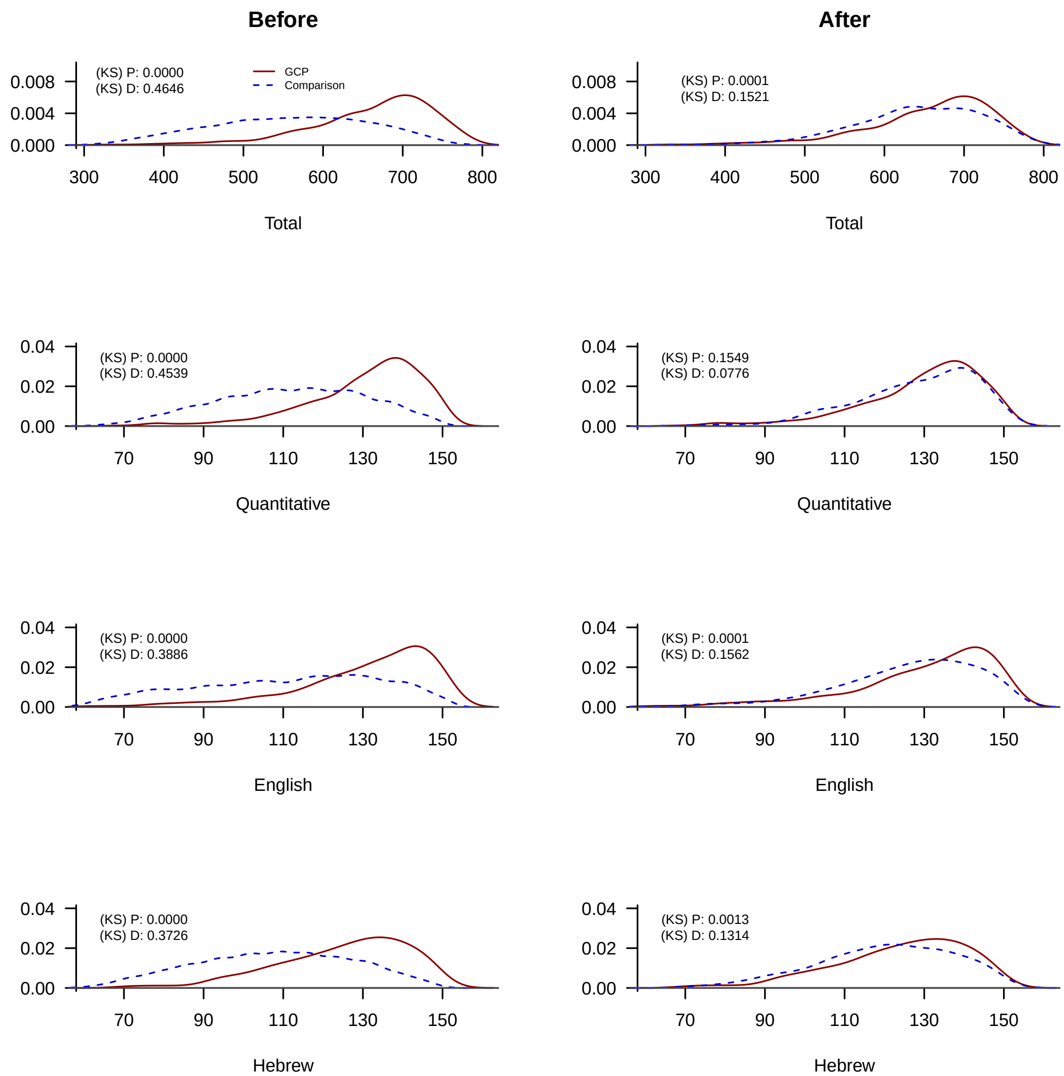
We begin by constructing a treatment indicator to represent participation in any of the first eleven GCPs in Israel, as detailed in the main section of our study. These eleven GCPs were identified using the high school identifiers and class numbers found in our dataset. Out of these, eight GCPs were established by 1990, the earliest year included in our dataset, two GCPs were introduced in 1995, and one started in 1999. During this time, nine of the schools provided only one gifted class per cohort, while the remaining two schools exclusively offered gifted classes.

Beginning in the 2000s, the Israeli Ministry of Education began documenting class types, with specific markers for gifted classes. We utilize this information to pinpoint new GCPs introduced during the timeframe of our study. However, our analysis deliberately excludes these newer GCPs to maintain a consistent treatment group, focusing solely on students from the earliest established gifted classes. Nonetheless, we leverage this data to remove any locality-cohort combinations that have access to a GCP from our comparison group, ensuring our control group does not include localities with available gifted education options.

A.3 Missing Values Imputation

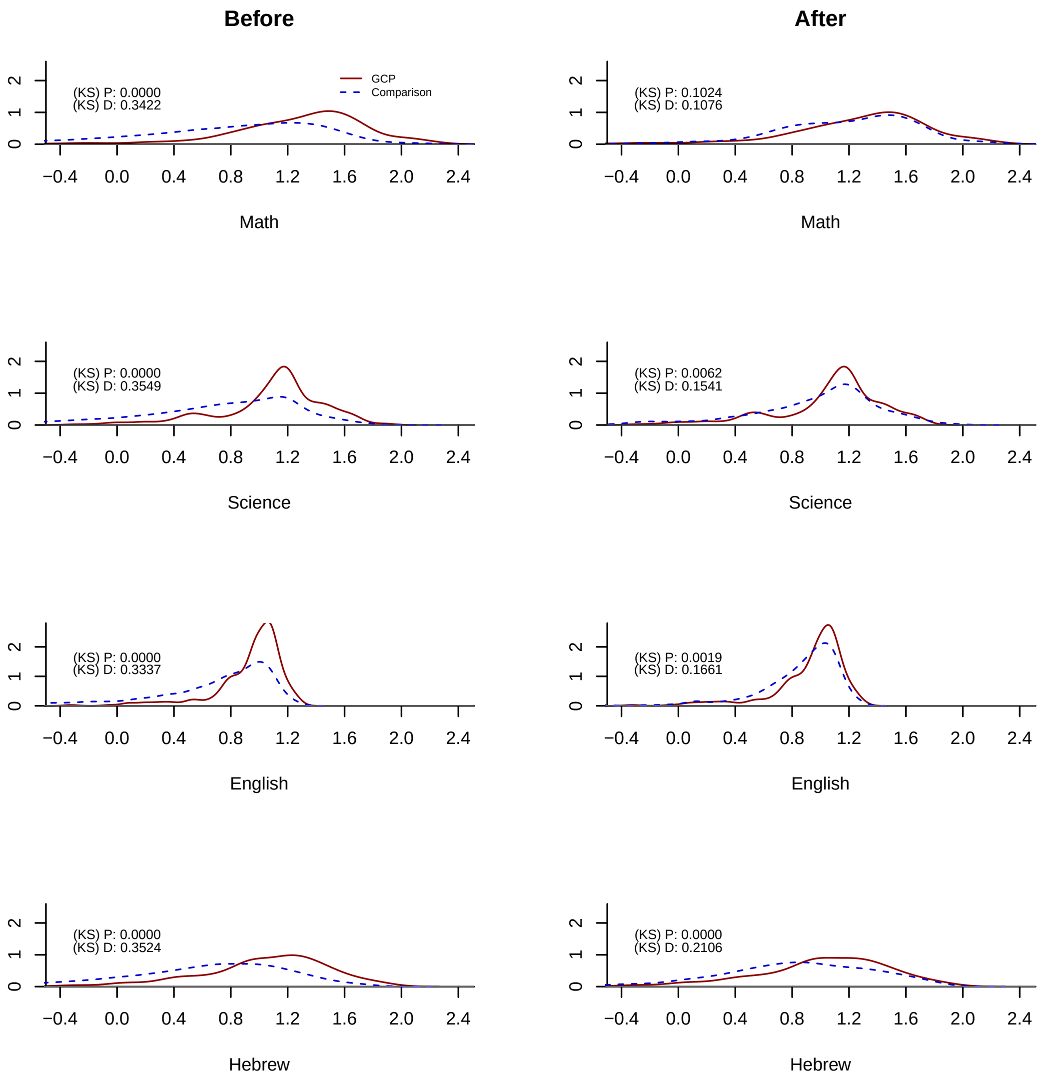
For some variables used in our matching specification, including parental years of education, our data had a small share of missing values. To handle this, we imputed the median value of this outcome in our sample and added an indicator variable for imputed observations. Then, we include in our matching specifications the variable itself, with the imputed data and the indicator for imputed observations.

Appendix Figure A1: UPET Scores, Before and After Matching



Notes: This figure plots the UPET scores distribution. The solid red lines represent the sample of GCP students, and the blue dashed lines represent the matched comparison group (which includes non-GCP students from other cities). The figures on the left show the distributions before the matching, and those on the right show the distributions after the matching. The figures also show the Kolmogorov–Smirnov test for the equality of the probability distributions. The sample includes students who participated in the 8th-grade Meitzav tests, about half of the students in cohorts of high-school graduates in 2006-2010. Note that the UPET scores are not included in the matching specification.

Appendix Figure A2: Middle-school Test Scores, Before and After Matching Using the Quantitative UPET Scores as the Ability Measure



Notes: This figure plots the pre-treatment 8th-grade Meitzav test scores distribution by groups. The solid red lines represent the sample of GCP students, and the blue dashed lines represent the matched comparison group (which includes non-GCP students from other cities). The figures on the left show the distributions before the matching and those on the right show the distributions after the matching (using quantitative UPET scores as the ability measure). The figures also show the Kolmogorov–Smirnov test for the equality of the probability distributions. The sample includes students from cohorts of high-school graduates in 2006-2010, who took the UPET until the age of 17. Note that the Meitzav test scores are not included in the matching specification.

Appendix Table A1: Standard Errors Calculation Methods

	No controls		With controls	
	Main	Abadie	Main	Bootstrap
	(1)	(2)	(3)	(4)
A. Matriculation				
Mean Matriculation Score	0.46	0.73	0.36	0.60
B. Higher Education				
% BA Attainment	2.41	2.43	2.35	2.05
% BA, Double Major	2.20	2.39	2.23	2.28
% BA, Math, CS, Stats	2.06	2.22	2.02	2.99
% BA, Engineering	2.14	2.42	2.15	2.98
% MA Attainment	2.16	2.57	2.20	2.52
C. Labor-market				
% Earnings Rank	1.58	1.72	1.56	2.15
% Knowledge Economy	2.63	2.96	2.58	4.00
D. Personal				
% Married	2.61	3.07	2.52	2.56

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

Notes: This table presents a validation test for the standard error calculation. Columns (1) and (2) compare the standard errors for the unconditional difference in the outcomes with the correction offered by Abadie and Imbens (2008). Columns (3) and (4) compare the standard errors for the conditional outcome difference with the clustered bootstrapped standard errors at the school level. The sample used for calculating the standard errors is our main sample, which includes about half of the students in cohorts of high-school graduates in 2006-2010 (those who participated in 8th-grade Meitzav tests).

Appendix Table A2: Comparison of Localities with and without GCPs

	Comparison	GCP	Estimated Effect
	(1)	(2)	(3)
Number of Students	800.81	3441.53	2640.72***
			(127.37)
% Males	48.85	48.49	-0.37
			(0.23)
Father Education	13.44	13.56	0.12**
			(0.05)
Mother Education	13.54	13.39	-0.15***
			(0.05)
% 5 Credits in Physics	9.12	8.30	-0.82***
			(0.18)
UPET Score	561.33	561.87	0.53
			(1.67)
% BA	46.32	44.38	-1.94***
			(0.63)
% MA	8.58	8.99	0.41**
			(0.17)
Students	555	555	1,110
Schools	144	11	155

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

Notes: The table shows the difference between localities with and without GCPs. The sample includes only students who participated in 8th-grade Meitzav tests, about half of the students in cohorts of high-school graduates in 2006-2010. Columns (1) and (2) show the means in locality-level outcomes for the matched comparison group and the treatment group (GCP participants). Column (3) shows the conditional difference (τ from equation (2)) in the outcome and its standard error (clustered at the school level).

Appendix Table A3: The Impact of GCP on Matriculation Test Scores

	Comparison	GCP	Estimated Effect
	(1)	(2)	(3)
Mean Composite Score	87.30	87.11	-0.19
			(0.33)
Math	87.81	84.11	-3.71***
			(0.85)
Hebrew	86.77	86.12	-0.64
			(0.40)
English	89.79	91.09	1.30***
			(0.44)
Bible	86.28	85.89	-0.40
			(0.56)
History	84.06	84.20	0.14
			(0.69)
Literature	81.28	80.90	-0.38
			(0.81)
Citizenship	84.12	83.33	-0.79
			(0.68)
Students	555	555	1,110
Schools	144	11	155

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

Notes: The table shows estimates for the impact of GCP participation on matriculation test scores. The sample includes only students who participated in 8th-grade Meitzav tests, about half of the students in cohorts of high-school graduates in 2006-2010. Test scores are measured on a 0-100 scale. Columns (1) and (2) show the means among the matched comparison group and the treatment group (GCP participants). Column (3) shows the conditional difference (τ from equation (2)) in the outcome and its standard error (clustered at the school level).

Appendix Table A4: The Impact of GCP on Labor-market Outcomes in 2018–2020

	Comparison	GCP	Estimated Effect
	(1)	(2)	(3)
A. Annual Earnings			
Earnings (100K NIS)	1.37	1.37	-0.00
			(8.49)
Log Earnings	11.39	11.40	0.01
			(0.08)
% Earnings Rank	54.59	54.16	-0.44
			(1.80)
% Top 10% Earners	0.20	0.18	-0.02
			(0.02)
B. Employment			
% Salaried Employment	88.29	86.94	-1.35
			(1.98)
% Self Employment	9.19	9.44	0.25
			(1.77)
C. Employment in Knowledge Producing Sectors			
% Knowledge Economy	50.45	51.50	1.05
			(2.90)
% Tech Services	35.50	38.87	3.37
			(2.81)
% Tech Manufacturing	9.01	4.11	-4.90***
			(1.52)
% Academic	10.27	11.68	1.40
			(1.87)
Students	555	555	1,110
Schools	144	11	155

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

Notes: The table shows estimates for the impact of GCP participation on labor-market outcomes during 2018–2020. The sample includes only students who participated in 8th-grade Meitzav tests, about half of the students in cohorts of high-school graduates in 2006-2010. Columns (1) and (2) show the means among the matched comparison group and the treatment group (GCP participants). Column (3) shows the conditional difference (τ from equation (2)) in the outcome and its standard error (clustered at the school level).

Appendix Table A5: Placebo Exercise, The "Impact" of Regular Classes on Outcomes, High-Achievers

	Comparison	GCP	Estimated Effect
	(1)	(2)	(3)
A. Matriculation			
Mean Composite Score	86.56	86.78	0.23
			(0.37)
B. Higher Education			
% BA Attainment	78.14	75.26	-2.88
			(2.47)
% BA, Double Major	15.67	17.20	1.53
			(2.45)
% BA, Math, CS, Stats	12.78	17.25	4.46**
			(2.22)
% BA, Engineering	24.12	20.95	-3.17
			(2.62)
% MA Attainment	20.21	19.33	-0.88
			(2.49)
C. Labor Market			
% Earnings Rank	58.18	60.23	2.05
			(1.80)
% Knowledge Economy	35.46	39.33	3.87
			(2.98)
D. Personal			
% Married	39.18	40.90	1.72
			(3.12)
Partner UPET Score	583.07	599.11	16.04
			(11.33)
Students	485	485	970
Schools	70	140	210

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

Notes: The table shows the results of a placebo exercise, estimating the effects of studying in regular classes (in high schools with no GCP, located in localities with a GCP), using the same matching algorithm we use throughout our analysis. The sample includes only students who participated in 8th-grade Meitzav tests, about half of the students in cohorts of high-school graduates in 2006-2010. Columns (1) and (2) show the means among the matched comparison group (students in other localities) and the treatment group (high-achieving students from regular classes in localities with a GCP). Column (3) shows the conditional difference in the outcome (τ from equation (2)) and its standard error (clustered at the school level). We find null results, as expected, supporting the validity of our design. GCP stands for Gifted Children's Program, and UPET stands for University Psychometric Entrance Test.

Appendix Table A6: Robustness of the Results, Alternative Matching Specifications

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
A. Matriculation							
Mean Composite Score	-0.19	-0.22	-0.12	-0.19	0.45	-0.33	-0.14
	(0.33)	(0.32)	(0.34)	(0.33)	(0.35)	(0.33)	(0.34)
B. Higher Education							
% BA Attainment	3.29	3.38	3.26	3.29	5.63**	-0.06	1.50
	(2.33)	(2.26)	(2.40)	(2.33)	(2.43)	(2.39)	(2.46)
% BA, Double Major	5.84**	6.08**	5.99**	5.84**	10.17***	4.28*	4.93*
	(2.51)	(2.46)	(2.56)	(2.51)	(2.40)	(2.54)	(2.54)
% BA, Math, CS, Stats	4.80**	4.44**	5.00**	4.80**	5.21**	2.64	1.55
	(2.22)	(2.18)	(2.26)	(2.22)	(2.29)	(2.26)	(2.33)
% BA, Engineering	-6.05**	-5.15**	-6.47**	-6.05**	-9.65***	-4.82*	-4.39*
	(2.38)	(2.34)	(2.42)	(2.38)	(2.46)	(2.42)	(2.42)
% MA Attainment	9.05***	9.78***	9.88***	9.05***	10.86***	8.72***	9.75***
	(2.55)	(2.50)	(2.62)	(2.55)	(2.52)	(2.57)	(2.70)
C. Labor-market							
% Earnings Rank	-0.55	-0.65	-0.41	-0.55	0.15	-2.02	-2.31
	(1.76)	(1.73)	(1.80)	(1.76)	(1.80)	(1.82)	(1.82)
% Knowledge Economy	-0.07	0.40	0.34	-0.07	-2.41	-3.39	-0.50
	(2.86)	(2.81)	(2.93)	(2.86)	(2.90)	(2.94)	(2.99)
D. Personal							
% Married	2.54	1.48	2.34	2.54	2.02	1.18	-4.11
	(2.78)	(2.73)	(2.85)	(2.78)	(2.83)	(2.90)	(2.99)
Partner UPET Score	22.10*	22.34*	21.34*	22.10*	22.44*	36.65***	17.80*
	(12.04)	(11.20)	(12.36)	(12.04)	(11.14)	(11.29)	(10.10)
Replacement				+			
Caliper	0.1	0.2	0.05	0.1	0.05	0.1	0.1
Estimation	Logit	Logit	Logit	Logit	XGB	Logit	Logit
Comparison localities	Other	Other	Other	Other	Other	Large	Same
Number of observations	1,110	1,156	1,060	1,110	1,156	1,058	1,008

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

Notes: The table shows estimates for the impact of GCP participation on outcomes, using different matching specifications: with/without replacement, changing the caliper, the method for estimating the propensity score, and the comparison group's pool. The sample includes only students who participated in 8th-grade Meitzav tests, about half of the students in cohorts of high-school graduates in 2006-2010. Each column shows the conditional difference (τ from equation (2)) in the outcome mentioned on the left and its standard error (clustered at the school level).

Appendix Table A7: UPET Scores by the Age of Taking the Test

	Total		Quantitative	
Coefficient	(1)	(2)	(3)	(4)
Intercept	546.61***	660.82***	113.21***	131.79***
	(0.83)	(2.86)	(0.15)	(0.54)
I(Age < 17)	-14.22	29.97	-1.48	8.01
	(10.45)	(33.00)	(1.96)	(6.21)
I(Age = 18)	1.92	8.25	-0.53**	0.93
	(1.42)	(5.46)	(0.27)	(1.02)
I(19 ≥ Age ≥ 21)	1.48	11.26**	-3.94***	-1.36
	(1.04)	(5.36)	(0.19)	(1.01)
I(Age > 21)	23.76***	26.17***	-0.48***	0.81
	(0.99)	(5.43)	(0.19)	(1.02)
Sample	All	GCP	All	GCP
Number of observations	83,698	1,431	83,698	1,431

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

Notes: This table presents the results of the regressions to predict students' UPET scores. The outcome variable in columns (1) and (2) is the total UPET score, and in columns (3) and (3) is the quantitative UPET score of each individual. The explanatory variables are indicator variables for ages at the test (presented on the leftmost common). The baseline sample includes students who participated in the UPET from cohorts of high-school graduates in 2006-2010. The sample in columns (2) and (4) is restricted to GCP participants. Standard errors are shown in parentheses.