Gifted Children Programs' Short and Long-Term Impact: Higher Education, Earnings, and the Knowledge-Economy

Preliminary and Incomplete

Victor Lavy, Yoav Goldstein*

June 13, 2023

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 significantly impacts university outcomes, such as choice of field of study and graduating with double majors, including double STEM fields. 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.

Keywords: Gifted children, higher education, knowledge economy

JEL Classification: I21 I23, I26, J24

*Victor Lavy: University of Warwick, Hebrew University, and NBER, V.Lavy@warwick.ac.uk. Yoav Goldstein: Berglas School of Economics, Tel Aviv University, yoavg2@mail.tau.ac.il. 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.

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 10^{th} grade (or 7^{th} 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 group 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

[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).

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 students from larger localities without a GCP yields the same results. Third, we show that the results are very similar when we select a comparison group from the same localities (with GCP). Thus, unobserved systematic differences between localities with and without a GCP are unlikely to bias our results. Finally, we also run a placebo exercise that estimates the effect of studying in regular classes (in different high schools in localities with a GCP). We find desired null results using the same matching algorithm to select a comparison group for these students, supporting the validity of our design.

We measure treatment effects 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 students gain undergraduate degrees, as almost all gifted children achieve this degree (above 90 percent). Still, we find a positive and large impact on the choice of field of study, and on graduating with a double major, including double STEM majors.

Importantly, we also analyze how the GCP affects career outcomes. While all gifted children have much better results 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 (ages 33–46). These findings suggest that gifted students have successful careers, but the GCP has no significant contribution to their career paths.

Examining the impact on personal outcomes, we find that GCP positively affects the likelihood of living outside Israel in adulthood, perhaps to study or work abroad. We also find that the GCP does not affect the probability of marriage or having kids.

Still, it positively affects the "quality" of the spouse (e.g., measured by the partner's test scores), driven by the marriages of GCP participants with their classmates.

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.

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 students' math, reading, and writing test scores. The GCP they investigate puts gifted students together in classrooms with other high achievers and offers an enriched curriculum. This study finds little if any, test score gains for gifted students but significant and positive improvements for their high-achieving peers. Cohodes (2020) finds that participation in dedicated classes with an accelerated curriculum positively impacted test scores and long-term outcomes for fourth through sixth graders in Boston public schools. Redding and Grissom (2021) use within-school and within-student comparisons and find that participating in a gifted program in public primary schools is associated with small achievement gains.

Bui et al. (2014) examine the effect of GCP on math, reading, and language test scores of middle-school students 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.[2] Bhatt (2009) also looks at U.S. middle school students but uses an instrumental variable strategy that exploits differences in GCP admission rules between schools. She finds positive test score gains, but this may partly reflect the sorting of students to schools.

[2] In a comparable experiment, Davis et al. (2013) exploit a fuzzy RD design to estimate whether GCP can help public schools retain gifted students. Close to the admission cut-off, they find that gifted students are more likely to stay in public schools.

Booij et al. (2016) examine the effect of a gifted and talented secondary education program as an individualized pull-out program in a specific school in the Netherlands. Like earlier studies, they also use a fuzzy RD design to estimate the impact on those at the program's margin of acceptance. They show that participants obtain higher grades, follow a more science-intensive curriculum (most notably for girls), and report stronger beliefs about their academic abilities. They also find that these positive effects persist in university, where students choose more challenging fields of study with, on average, higher returns. In another paper, they analyze similar programs in additional schools to examine their effects on academic performance (Booij et al., 2017). They use different strategies to show that the effects are positive and stronger for students further away from the admission cutoff.

This paper makes several significant contributions to this literature. First, it is based on an experienced gifted children's program for over three decades in separate gifted classes or schools. We analyze a comprehensive set of outcomes in the medium and long run. Beyond completing university degrees, we examine the university choice of field studies, especially in STEM. Assessing whether the program affects career and family outcomes takes a longer-run perspective. Therefore, we follow GCP participants after completing their degrees and examine their earnings and family formation patterns. Uniquely, we examine GCP participants' contribution to the knowledge economy by looking at their integration into hi-tech and academic institutions. Secondly, an essential contribution of this paper is the analysis of treatment heterogeneity by giftedness level. This distinguishes this paper from most earlier studies that used RD designs to identify GCP effects, focusing more on the marginally eligible students for such programs. Another significant contribution is our distinction in estimating treatment effect by the length of GCP.

We relate some of our findings to theories and hypotheses in the literature in psychology about gifted children. It includes the literature regarding the affective and personality development of the socio-emotional characteristics of gifted children. The literature on 'big fish small pond' is perhaps key in understanding our finding of no effect of GCP 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 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, including more double STEM 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 the empirical methodology. Next, we present the results in section 5. Finally, section 6 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 (Boettger and Reid, 2015; Vrignaud et al., 2005; Mönks et al., 2005). 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 within 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. In parallel, 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.

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 (6^{th} 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

[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).

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

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 GCP 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 1992–2013. 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. Finally, we also analyze the outcomes of earlier and later cohorts as supplementary analyses.

2.3 Israel's High School and Higher Education Systems

When entering high school (10^{th} grade), students enroll in the academic or non-academic track. Students enrolled in the academic track receive a matriculation certificate (Bagrut) if they pass a series of national exams in core and elective subjects between 10^{th} and 12^{th} grade. Depending on difficulty, students choose to be tested at various proficiency levels, each awarding one to five credit units per subject. Advanced level subjects award students more credit units (5 relative to 4 for an intermediate level and 3 for a basic level); a minimum of 20 credit units must qualify for a matriculation certificate. Courses which award five credits are equivalent to advance placement courses in the US high school system.

[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.

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 Lavy (2020, 2021).

Israel's post-high school academic schooling system includes ten universities (one of which confers only graduate degrees) and 50 colleges that confer undergraduate degrees (some also give master's degrees).[6] 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.

[6] A 1993 reform sharply increased the supply of postsecondary schooling in Israel by creating publicly funded regional and professional colleges.

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 Metzav exams taken in four subjects (science, math, English, Hebrew) during primary school (5^{th} grade) and middle school (8^{th} grade). The second is the University Psychometric Entrance Test (UPET), which includes test scores in three domains (quantitative, verbal and English).[7]

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 8^{th}-grade (5^{th}-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 8^{th}-grade test scores as the proxy for ability. In 2020, the latest year in the labor market data we use (post-secondary education data is available for 2021, too), the youngest cohort in the main sample is 28 years old, 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 standardized 8^{th}-grade tests, about half of the students in each cohort. Before matching, it includes 620 GCP participants and 63,590 students from other localities, which are included in our comparison group pull (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 5^{th}-grade test scores. The

[7]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.

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. Again, 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

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. As a robustness test, we also calculate the standardized mean scores by standardizing all test scores at the cohort level.

Post-secondary education. We define outcomes for getting any BA degree and a BA degree from an Elite University.[8] 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.[9] We follow the grouping definition of the CBS (for more details, see Appendix A). 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, Industrial equipment for control and supervision, medical and scientific equipment, and Aircraft. *High-tech Services industries*: Telecommunica-

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

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

tions, Computer, and related services, Research and development. *Academic*: Colleges of education, other extensions of foreign universities, Universities. *Knowledge*: any of the above.[10]

Earnings. We focus on the total annual earnings from 2018–2020. We use the earnings rank conditional on age as the main outcome variable. We also analyze the natural log of the earnings and earnings without outliers.[11] 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. We also use indicators for being married or having the first child before age 28. 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 GCP programs' effect in the U.S. (Card and Giuliano, 2014; Bui et al., 2014). This design exploits the admission cut-off to GCP.[12] 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 achievement for infra-marginal gifted children. We chose gifted children from cities where GCP was not offered at the time as a comparison group.

[10] 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.

[11] We dropped from the sample all observations that are six or more standard deviations away from the mean to account for earnings data outliers. Significantly few observations are dropped from the sample each year, and the results are not qualitatively affected by this sample selection procedure.

[12] Abdulkadiroğlu et al. (2014) used the same identification strategy to estimate the effect of elite schools in Boston and NY.

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 the decision to enroll 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 Metzav exams taken in four subjects (science, math, English, Hebrew) during 8^{th} grade. We standardize all test scores at the cohort level.

Individuals select 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 10^{th} grade and most likely reflects the student's academic motivation and ambition at this stage. Since a student's study program is pre-determined, we can use it to match GCP participants to students in localities where GCP was not offered.

Additionally, individuals might select into GCPs according to their socioeconomic status and family backgrounds. Thus, we also use measures for parental education and age at birth, the number of siblings, family order, and the country of birth when choosing the comparison group.

4.2 Constructing the Comparison Group

We use the following two-step propensity score matching algorithm to construct a comparison group:[13] We start by estimating the propensity score equation using the

[13]Rosenbaum and Rubin (1983) proposed an approach that circumvents the curse of dimensionality when using selection on observables for the identification of causal effects. They provide proof that if treatment

sample of GCP participants and students from other localities (the comparison group pull):

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

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 8^{th}-grade test scores in each subject (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), the five most common elective subjects, and the five most common elective subjects among GCP participants relative to the general population. We also include the total number of matriculation credits (with a quadratic term). Finally, we include the following family background variables: The father's and mother's years of schooling (with quadratic terms), the mother's age at birth, 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 being 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).[14]

Figure 1 presents the propensity score distribution before and after the matching. We match 550 of the sample of 620 GCP participants. The unmatched are mainly from

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.

[14] We further validate that our main results are not sensitive to the specification by running alternative matching specifications. The results are shown in Appendix Table A5 and are identical to our main results.

the top of the propensity score distribution.[15] 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 8^{th}-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 detailed summary 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 in terms of parental characteristics. Interestingly, the groups are balanced on variables not included in the matching specification, such as parental earnings, parental country of birth, and paternal age at birth. For example, the father's (mother's) average yearly earnings in 2003 were 153,000 (87,000) NIS in the treated group and 159,000 (79,000) in the comparison group. These differences of less than 8,000 NIS yearly (about 2,000 USD) are not statistically different from zero. Interestingly, gifted children, either GCP participants or others, come from higher socioeconomic backgrounds than regular Israel 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 14.5).

Table 2 shows that both groups are balanced regarding student characteristics, with only 1 out of 10 estimates being statistically different than zero. 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

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

studies at very high rates (about 70% and 90%). GCP participants' most common elective subjects are 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.[16]

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).[17]

4.4 Validating the Identification Strategy

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 the locality. Therefore, we examine whether families with GCP participants had a higher mobility rate before 10^{th} grade than families with gifted children that did not

[16] We also validate that the standard errors are similar to the correction specified by Abadie and Imbens (2008) and to clustered bootstrapped standard errors.

[17] 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.

participate in a GCP. We find no such differential mobility rate. The rate of families who relocate between the 9^{th} and 10^{th} grades is 7.1% in the comparison group and 6% in the treatment group. The difference of 1.1p.p. is statistically insignificant ($p = 0.57$). We also examine relocation patterns at earlier ages and find no evidence that families of GCP participants relocate more.

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 similar characteristics, including parental background and academic outcomes, to those in other cities (see Appendix Table A3). Second, restricting the comparison group to include students from larger cities yields the same results. Third, we show that the results are very similar when we use a comparison group of students from the same cities (with a GCP). 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 into GCPs, and affects the outcomes that we analyze. Given that our data set includes very detailed information on each student's academic ability, academic motivation, and family background, we think it is unlikely such a variable could bias our results.

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 in localities with a GCP. 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. Table 3 shows that this exercise yields desired null results concerning the primary outcomes, further supporting our research design'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 using the main sample, including GCP participants, and matched comparison students who graduated high school between 2006–2010. In Section 5.5, we extend the analysis to additional samples to show that the results are robust, and in Section 5.6, we analyze the heterogeneity of the effects.

5.1 High School Achievement

We start by examining whether GCP participation affects 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 tiny and insignificant. This suggests that the GCP does not affect matriculation achievement on average. Additionally, Appendix Table A1 shows the estimates for the average GCP effects on test scores in all compulsory subjects. We find a negative effect on math test scores and positive effects on English, Hebrew, and literature.[18] However, all estimated effects are small in magnitude, suggesting that the GCP has no major effects on high school achievement.

The pattern of no major effects of GCP on test scores in matriculation exams is especially intriguing given the abundance of educational inputs that GCP participants enjoy relative to the comparison group we use. In Appendix Table A2, we compare the two groups class-level inputs. GCP classes are much smaller, the socio-economic background of peers in these classrooms is much higher, and the averages of outcomes are much better. We also based on details of GCP in Israel that their teachers are more

[18]Note that we report throughout the paper only the estimates that are based on Equation 2 including all control variables, but we also validate that the results are very similar if we exclude them. This is not a surprise, as we also show that the groups are very balanced in any important measure (in subsection 4.4).

qualified and receive additional training and that the budget per student is higher. 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 exams 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.[19] 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).[20] Although the BFLPE model is not specific to

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

[20]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

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).

5.2 Higher Education Outcomes

We proceed by examining the GCP effects on higher education outcomes. Table 4 presents the results. First, we show that GCP participants complete their undergraduate degrees at earlier ages (Panel A). This result is perhaps because many GCPs offer their participants to start university studies during high school. However, the difference becomes statistically insignificant when using an indicator for achieving a degree at any age. Panel B of the table shows that GCP participants tend to attain more advanced degrees, including those in elite universities. However, this could be because many GCP participants started their university studies earlier. Indeed, Panel C shows that a higher share of the matched comparison students are still enrolled in an

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

undergraduate degree.[21]

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 movement from engineering programs to physical sciences programs. The estimated effect is very large on the likelihood of studying engineering and Physical sciences. The first declines by 5p.p, amounting to a fall of 20 percent. The second increase by three p.p, implying a 60 percent rise. Panel B shows that the likelihood of achieving any non-STEM degree is not changing, too. Notably, the program significantly affects the likelihood of getting a double major. Panel C shows that the estimated effect is meaningful, 6p.p. relative to a baseline of 16.5%, implying a 35% relative increase.[22] There is also a considerable increase of about 80% in the probability of attaining a double STEM degree.

Graduating with a double major, especially with two STEM subjects, can be related to the multipotentiality of gifted children. This concept has been 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). Multi-potentiality is widely cited as a characteristic of the most gifted individuals who have 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. The realization of 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

[21]This implies that we should interpret the results in this section cautiously since not all students have finished university studies. However, in section 5.5, we show that the main results are similar when analyzing the older cohorts (who already completed their studies by 2020).

[22]This result is highly statistically significant, and it remains so even when applying a conservative Bonferroni correction for multiple hypothesis testing. The standard p-value for the increase in testing is 0.0045. Therefore, even if we take a very conservative approach of calculating the Bonferroni-corrected p-value for nine tests, we get a significant result ($p = 0.0405$).

that guide them? To what extent does the environment impact these choices? Studies in educational psychology on the formation of 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.

5.3 Labor Market Outcomes

An important question regarding gifted children's education is whether participation in a GCP significantly affects their career outcomes. We provide here 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 1% or 10% earner does not change. The estimated effects are -0.73p.p. and -0.93p.p. with standard errors of 1.35 and 2.49 (relative to baselines of 6% and 25%). Additionally, Panel B of Table 6 shows that the GCP has no significant effects on the likelihood of being employed or self-employed.

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 (full details in the Appendix A). Panel C shows that the likelihood of being employed in the knowledge economy is similar among GCP participants and comparison students. Note that this does not imply that gifted children do not contribute to the knowledge economy in Israel but that GCP does not affect the probability that gifted children will work in the knowledge economy.

5.4 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 evidence for a significant increase in the likelihood of living outside Israel in 2020. This may be driven by GCP participants who work or study abroad.[23] 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 UPET score of the partner. 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 matches within educational institutions. Our findings suggest that GCPs also matter for the marital matches of gifted children.

5.5 Results' Robustness and Persistence

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. Column (1) shows the results of our main specification for comparison. Columns (2)-(3) of Appendix Table A5 show that the results are very

[23] Importantly, we check, and the estimation of the GCP effects on labor-market outcomes do not change if we exclude those that live outside of Israel in 2020, as can be seen in Appendix Table A4.

similar when using other calipers (0.2SD and 0.05SD). 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. Instead of a logit model, we use a gradient-boosting algorithm (Chen and Guestrin, 2016) that allows a more flexible fit. We find very 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 where there is 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 above-median number of students in the locality) and from localities with a GCP (results are shown in columns (6)-(7) of Appendix Table A5).

Using 5^{th}-grade test scores as the ability measure. 5^{th}-grade Metzav test scores are available for high school graduates of 2009 or later. We use the graduates between 2009–2013 to validate that our main results are robust for using these test scores as the ability measure in the matching specification. Column (2) of Appendix Table A6 shows that the results are very similar to our main results. Note that the positive effect on the likelihood of achieving a BA degree is not surprising since these students are younger in 2020, and many are still enrolled in universities. Therefore, the only difference from our main results is the small positive effects on the mean composite matriculation score. While the effect is statistically significant, it is still very small in magnitude.

Using UPET scores as the ability measure. Remarkably, even though most students in Israel start college education at 22-23, most GCP participants take this test while in high school and even before midway into the program. Appendix Figure A1 presents the distributions of the UPET scores (total score, quantitative reasoning, verbal reasoning, and English) before and after the matching (using our primary sample and specification). The UPET score distributions show a clear advantage for GCP participants before the matching. However, after the matching, most of the differences in test scores are eliminated, and the quantitative score distribution is even statistically indistinguishable.

We also show evidence that the UPET scores are not affected by the age of taking the test, further supporting the idea that they measure intelligence, which is not majorly affected by life experiences. Appendix Table A7 shows that the differences in test scores between students who took their tests at early and later ages are relatively small. This is especially true for GCP participants and their quantitative scores (column (4) of the table), where all differences are insignificant.

Perhaps the finding that UPET test scores do not vary among highly talented individuals by the testing age is not surprising because its structure and content are very similar to the SAT and CAT used in the U.S. for the same purpose. In addition, these tests were shown to be highly correlated with I.Q. and other ability measures (Koenig et al., 2008; Beaujean et al., 2006; Frey and Detterman, 2004), and we should not expect them to vary much by age. So is the UPET, but the evidence for its correlation with IQ test scores is more limited.

Importantly, this finding allows us to view the UPET score as pre-determined and to use it in our matching specification. We also show that the results are similar when using UPET scores instead of 8^{th}-grade Metzav test scores. We focus on the sample of early takers (who took the UPET until 17) and use only the quantitative UPET score as the ability measure. Column (3) of Appendix Table A6 shows that the results are identical to the main results. These findings are important as they allow us to analyze the GCP effects for earlier (older) cohorts. For them, the only ability measure available in our data set is the UPET scores.

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. Columns (4)-(6) of Appendix Table A6 shows that the main results are very similar when excluding some or all of these variables. The only difference from our main results is the small positive effects on the mean composite matriculation score, which is small in magnitude. The strong positive effect on the likelihood of a double major degree and the nill effects on labor-market outcomes remain identical.

Analyzing the GCP effects for older cohorts, using UPET scores as the ability measure. We use the sample of high school graduates between 1992 and 2005 to examine the persistence of the results in the longer term (these individuals were ages 33-46 in 2020). We start by restricting the sample to those who took their UPET early and use only quantitive UPET scores as the ability measure in the matching specification. Column (1) of Appendix Table A8 shows that the results persist, with strong positive effects on the likelihood of achieving a BA degree and no effects on labor-market outcomes such as earnings or being employed in the knowledge-producing sectors.

Additionally, Appendix Figure A2 compares the earnings distributions of both groups of gifted students. Similar to what we found in the primary sample, there are no differences in earnings between GCP participants and the matched comparison students. 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.

We also show that the results are similar when extending the sample to all UPET test takers and using the UPET scores in all three domains (Columns (2)-(4)).

5.6 Heterogeneity of the GCP Effects

In this subsection, 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 the extent that the effects vary along z. We report the results regarding the heterogeneity of the effects on the main outcomes, but we have also checked other outcomes, and the general impression is similar.

Gender heterogeneity. First, we examine whether the GCP affects boys and girls differently. Column (1) of Table 8 shows the estimated differences in the effects of GCP on boys and girls for the main outcomes. Appendix Table A9 provides more details, including the estimated effect and the baseline mean in the outcome for each group. We do not detect any significant difference, suggesting that the GCP affects boys and girls similarly. One caveat for this conclusion is the relatively small sample size, implying that we could only detect large differences. Therefore, we also estimate equation 3 using the sample of older cohorts. The results are reported in Column (1) of Appendix Table A13. We find that for these cohorts, the positive effects on double major degrees are driven by the girls.

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 father education of 15 years (the minimal number of years required to attain a BA degree) or more. Column (2) of Table 8 shows the estimated differences (Appendix Table A10 provides more details). Overall, GCP effects do not vary by student's SES background.[24] 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.

Level of 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. The results are shown in Col-

[24] We do find a stronger effect on BA degrees for high-SES students. However, it may be due to some students still studying towards a degree. Indeed, when analyzing the outcomes of the older cohorts, we do not find any significant differences (Column (2) of Appendix Table A13).

umn (3) of Table 8, and in Appendix Table A11. We do not detect any significant difference in the effects.[25]

Length of participation in the GCP. Half of the GCP participants of the main sample participated in a GCP since middle school, and the rest participated in a GCP only during high school. This setting allows us to address an important dimension of the GCP treatment effects heterogeneity, namely whether participation in a GCP for a more extended period (7th-12th grades) affects student outcomes differently from shorter participation (10th-12th grades). Interestingly, we find no evidence for heterogeneity in this dimension (Column (4) of Table 8, and in Appendix Table A12), suggesting that GCP effects do not vary by the length of participation in the program.

6 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.

[25] When analyzing the older cohorts, we do find one significant difference regarding the effects on a double major degree (Column (3) of Appendix Table A13).

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 negative effects on academic performance). Second, GCP’s studies program incorporates advanced concepts and topics not directly relevant to the standard curriculum. Thus, GCP participants may get educational benefits not manifested in higher matriculation test scores.

Secondly, we find a large and significant effect on a double major, including mainly two STEM subjects. This effect may reveal the multipotentiality of gifted children and their difficulty selecting one area of interest to focus. The focus on prestigious and highly regarded fields of study, such as physical sciences, is consistent with the view that gifted children are under social pressure by parents and social circles to ‘maximize’ their potential and not to ‘waste’ it on areas that are not too challenging intellectually. As a result, we should not be surprised by our findings of no effect on earnings in adulthood as the career path of gifted children is not necessarily guided by consideration of maximizing the financial return to their ability. Perhaps surprising is GCP’s ‘no’ effect on integrating gifted children into work in sectors that produce ‘new’ knowledge. One explanation could be that gifted children are directed into these industries in advance. Thus, the GCP plays no important role in these decisions.

Against the benefit 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.

Abdulkadiroğlu, Atila, Joshua Angrist, and Parag Pathak, “The elite illusion: Achievement effects at Boston and New York exam schools,” *Econometrica*, 2014, *82* (1), 137–196.

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.

Bhatt, Rachana R, “The impacts of gifted and talented education,” *Andrew Young School of Policy Studies Research Paper Series*, 2009, (09-11).

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.

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.

Davis, Billie, John Engberg, Dennis Epple, Holger Sieg, and Ron Zimmer, "Bounding the impact of a gifted program on student retention using a modified regression discontinuity design," *Annals of Economics and Statistics/ANNALES D'ÉCONOMIE ET DE STATISTIQUE*, 2013, pp. 10–34.

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, 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.

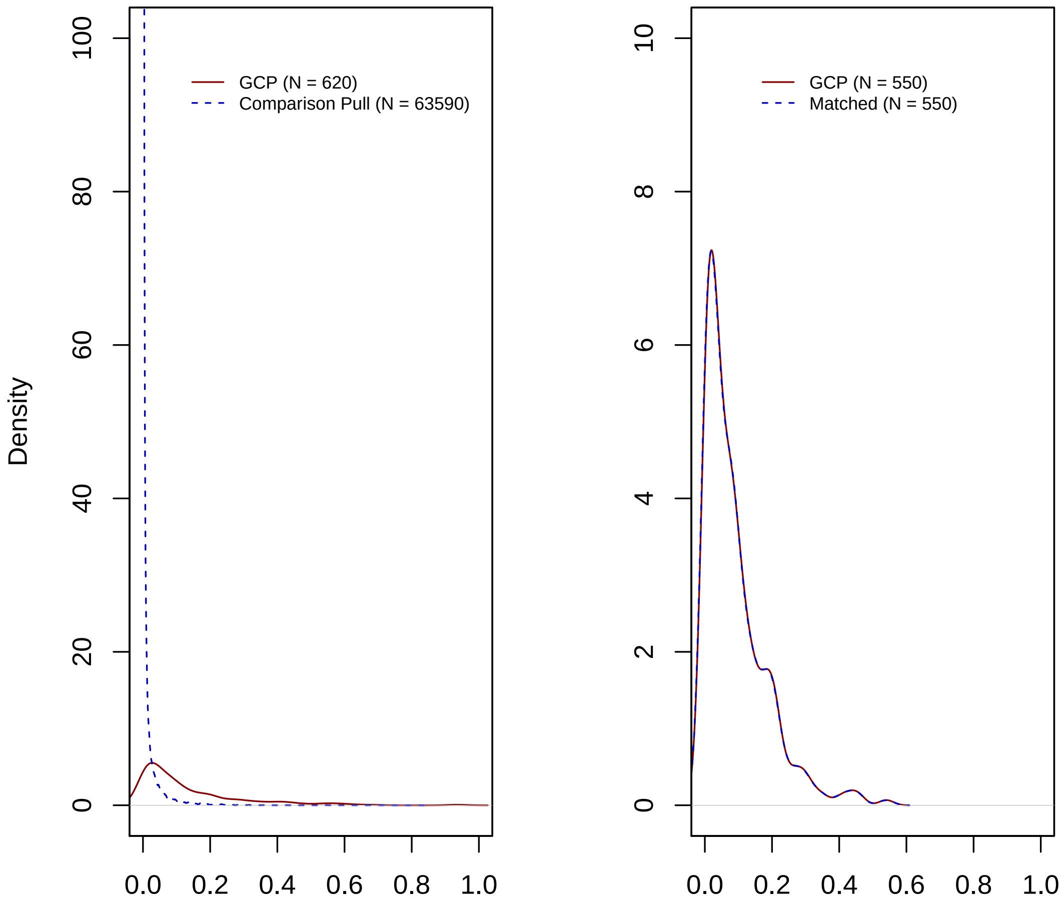
Vrignaud, Pierre, Denis Bonora, and Annie Dreux, "Counselling the gifted and talented in France: minimizing gift and maximizing talent," *International Journal for the advancement of counselling*, 2005, *27*, 211–228.

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.

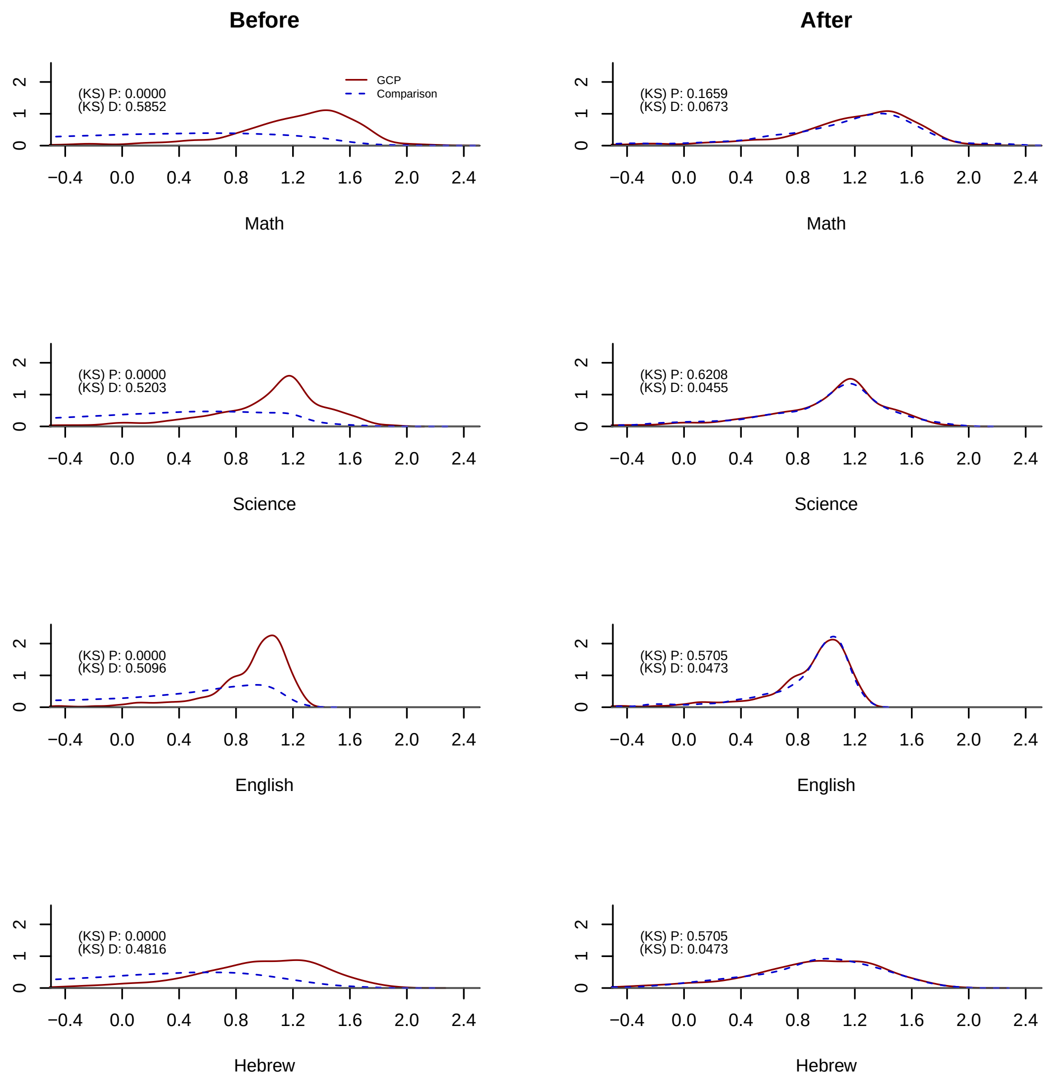
_ , _ , Gerald Matthews, and Richard D Roberts, "Assessing emotional intelligence in gifted and non-gifted high school students: Outcomes depend on the measure," *Intelligence*, 2005, *33* (4), 369–391.

Figure 1: Propensity Score Distributions, Before and After Matching



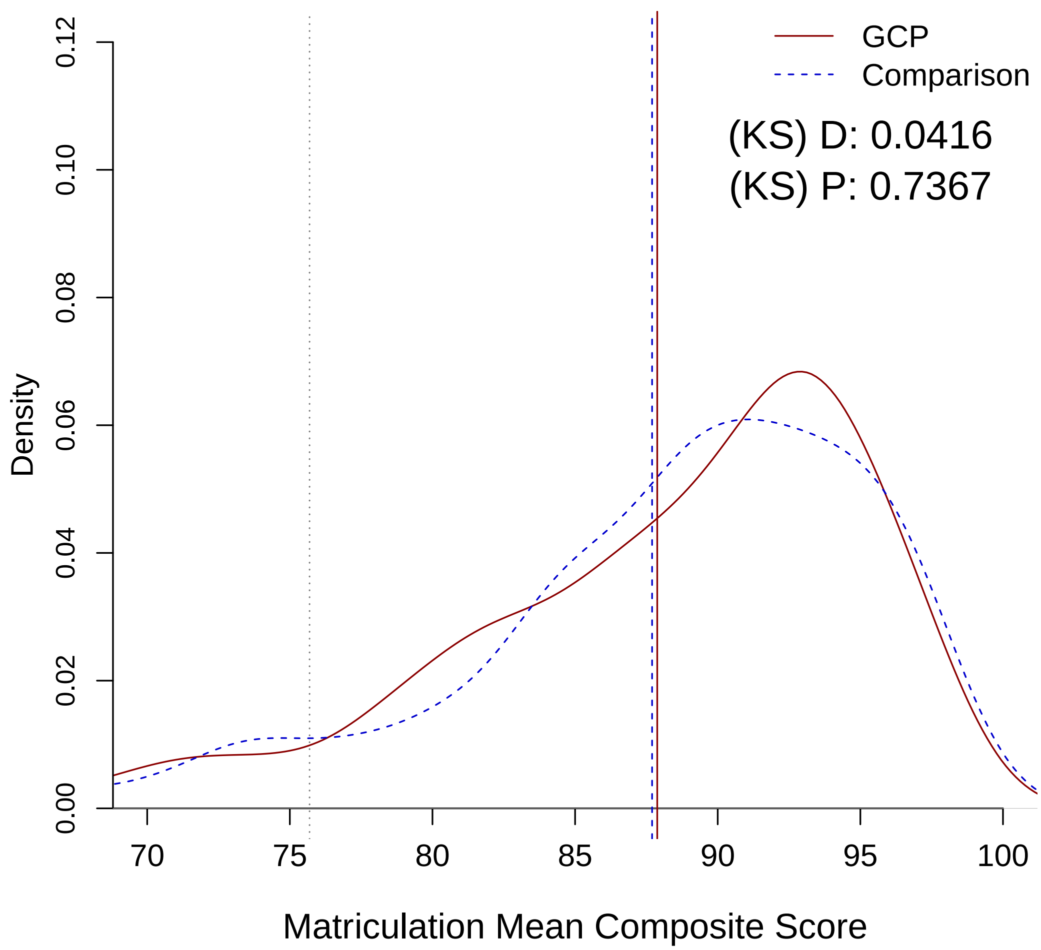
Notes: This figure plots the propensity score distribution by groups. The solid red line represents the sample of GCP students, and the blue dashed line represents the comparison group (which includes non-GCP students from other cities). The graph on the left shows the distributions before the matching, and the graph on the right shows the distributions after the matching. The sample includes only students who participated in the Metzav middle school test during their 8^{th} grade, about half of the students in cohorts of high-school graduates in 2006-2010.

Figure 2: Pre-treatment Middle-school Test Scores, Before and After Matching



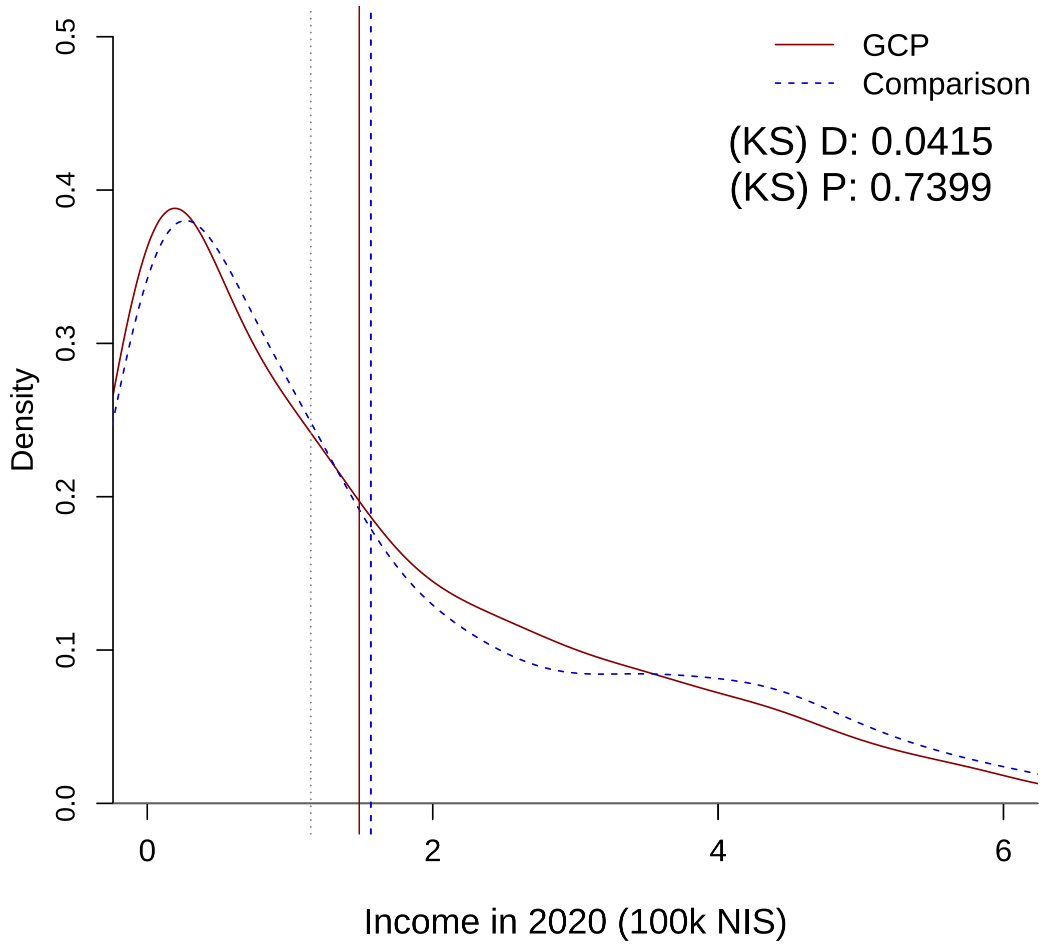
Notes: This figure plots the pre-treatment 8^{th}-grade Metzav test scores distribution by groups. The solid red line represents the sample of GCP students, and the blue dashed line represents the comparison group (which includes non-GCP students from other cities). The graphs on the left show the distributions before the matching, and those on the right show the distributions after the matching. The graphs also show the Kolmogorov–Smirnov test for the equality of the probability distributions. The sample includes students who participated in the 8^{th}-grade Metzav tests, about half of the students in cohorts of high-school graduates in 2006-2010.

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



Notes: This figure plots the mean composite matriculation (Bagrut) scores 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 graphs report the Kolmogorov–Smirnov test for the equality of the probability distributions. The vertical lines represent the averages, and the difference between the averages is 0.18 (insignificant with a p-value of 0.72). The dotted grey line represents the comparison group's pull, including all students from cities with no GCPs. The sample includes students who participated in the 8^{th}-grade Metzav tests, about half of the students in cohorts of high-school graduates in 2006-2010.

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 graphs report the Kolmogorov–Smirnov test for the equality of the probability distributions. The vertical lines represent the averages, and the difference between the averages is -0.08 (insignificant with a p-value of 0.42). The dotted grey line represents the comparison group's pull, including all students from cities with no GCPs. The sample includes students who participated in the 8^{th}-grade Metzav tests, about half of the students in cohorts of high-school graduates in 2006-2010.

Table 1: Balancing Table, Parental Characteristics

	Comparison	GCP	Difference
	(1)	(2)	(3)
A. Earnings, 2003			
Father	1.59	1.53	-0.06
			(0.11)
Mother	0.79	0.87	0.08
			(0.06)
B. Earnings, 2006			
Father	1.99	1.76	-0.23
			(0.21)
Mother	0.95	1.02	0.07
			(0.07)
C. Years of Education			
Father	15.10	15.11	0.01
			(0.18)
Mother	14.99	15.07	0.08
			(0.17)
D. Born in Israel			
Father	85.27	82.73	-2.55
			(2.21)
Mother	66.73	64.18	-2.55
			(2.87)
E. Age at Birth			
Father	32.56	32.67	0.11
			(0.34)
Mother	29.16	29.59	0.42
			(0.31)
Students	550	550	1,100
Schools	138	11	149

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

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 8^{th}-grade Metzav tests, about half of the students in cohorts of high-school graduates in 2006-2010. Earnings (panels A and B) are measured in 100K NIS. Years of education (panel C) are measured in years. The outcomes in panel D are indicator variables for Israel being the country of birth, and it is measured in percentages (0-100). Age at birth (panel E) is measured in years. 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 unconditional difference in the outcome and its standard error. Among the variables in the table, only parental years of education and maternal age at birth are included in our matching specification.

Table 2: Balancing Table, Student Characteristics

	Comparison	GCP	Difference
	(1)	(2)	(3)
A. Background			
Number of Siblings	1.69	1.70	0.01
			(0.08)
Family Order	1.78	1.77	0.00
			(0.06)
Born in Israel	85.27	82.73	-2.55
			(2.21)
B. Matriculation			
Total Credits	28.50	29.57	1.06**
			(0.41)
Math, 5 Credits	71.45	75.45	4.00
			(2.66)
English, 5 Credits	88.73	91.45	2.73
			(1.80)
Physics, 5 Credits	46.91	47.82	0.91
			(3.01)
CS, 5 Credits	47.64	48.00	0.36
			(3.01)
Chemistry, 5 Credits	32.36	28.55	-3.82
			(2.78)
Biology, 5 Credits	14.73	15.64	0.91
			(2.17)
Students	550	550	1,100
Schools	138	11	149

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

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 8^{th}-grade Metzav tests, about half of the students in cohorts of high-school graduates in 2006-2010. The indicator variables, for being born in Israel and achieving five credits in different matriculation subjects, are measured in percentages (0-100). 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 unconditional difference in the outcome and its standard error. All the variables in the table are included in our matching specification.

Table 3: Placebo Exercise, The "Impact" of Regular Classes on Outcomes

	Comparison	GCP	Estimated Effect
	(1)	(2)	(3)
A. Matriculation			
Mean Composite Score	75.84	76.31	-0.03
			(0.47)
B. Post-secondary Education			
BA Attainment	47.40	46.40	-2.94
			(2.35)
BA, Double Major	7.37	6.37	-1.30
			(1.45)
C. Labor-Market, 2020			
Annual Earnings	1.01	1.01	-0.04
			(0.06)
Knowledge Economy	17.09	17.92	-0.31
			(2.13)
D. Personal, 2020			
Out of Israel	4.86	4.52	-0.36
			(1.27)
Students	594	594	1,194
Schools	157	157	314

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

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 8^{th}-grade Metzav 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 (students in 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.

Table 4: The Impact of GCP on Post-secondary Education Outcomes

	Comparison	GCP	Estimated Effect
	(1)	(2)	(3)
A. BA Degrees			
Until Age 21	0.91	5.82	4.93***
			(1.08)
Until Age 25	13.27	23.27	9.14***
			(2.29)
Ever	75.64	80.00	1.39
			(2.37)
Elite University	32.73	38.91	4.21
			(2.75)
B. MA Degrees			
Any	16.18	28.36	11.39***
			(2.52)
Elite University	6.00	10.00	3.54**
			(1.61)
C. Enrollment in 2020			
Any Degree	26.36	27.64	0.42
			(2.68)
BA	12.91	7.99	-4.92**
			(1.88)
Students	550	550	1,100
Schools	138	11	149

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

Notes: The table shows estimates for the impact of GCP participation on post-secondary education outcomes. The sample includes only students who participated in 8^{th}-grade Metzav tests, about half of the students in cohorts of high-school graduates in 2006-2010. All outcomes are measured in percentages (0-100). 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).

Table 5: The Impact of GCP on University Fields of Study

	Comparison	GCP	Estimated Effect
	(1)	(2)	(3)
A. STEM Fields			
Any STEM	46.18	45.09	-1.88
			(2.73)
Math, CS, Statistics	17.09	20.00	2.94
			(2.25)
Engineering	22.55	18.00	-4.86**
			(2.38)
Physical Sciences	4.91	8.36	3.16**
			(1.48)
Biological Sciences	4.36	4.55	-0.06
			(1.23)
B. Non STEM Fields			
Any Non STEM	29.45	32.73	3.27
			(2.59)
C. Double Majors			
Any Double Major	16.55	24.18	6.82***
			(2.40)
Double STEM	4.73	8.73	3.94**
			(1.47)
STEM & Other	3.27	2.91	-0.43
			(1.04)
Students	550	550	1,100
Schools	138	11	149

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

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 8^{th}-grade Metzav tests, about half of the students in cohorts of high-school graduates in 2006-2010. All outcomes are measured in percentages (0-100). 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).

Table 6: The Impact of GCP on Labor-market Outcomes in 2020

	Comparison	GCP	Estimated Effect
	(1)	(2)	(3)
A. Annual Earnings			
Earnings (100K NIS)	1.65	1.60	-0.06
			(0.10)
Log Earnings	11.60	11.64	0.05
			(0.09)
Earnings Rank	58.42	58.03	-0.64
			(1.76)
B. Employment			
Salaried Employment	82.36	78.73	-3.91
			(2.41)
Self Employment	6.73	4.55	-2.04
			(1.37)
C. Employment in Knowledge Producing Sectors			
Any Knowledge Sector	40.91	39.82	-1.61
			(2.89)
Tech Services	31.45	31.64	-0.30
			(2.72)
Tech Manufacturing	2.18	1.82	-0.21
			(0.90)
Academic	7.27	6.36	-1.09
			(1.60)
Students	550	550	1,100
Schools	138	11	149

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

Notes: The table shows estimates for the impact of GCP participation on labor-market outcomes. The sample includes only students who participated in 8^{th}-grade Metzav tests, about half of the students in cohorts of high-school graduates in 2006-2010. In panel A, earnings are measured in 100K NIS, the log is measured in log points, and the rank and top earners are measured in percentages (0-100). Employment outcomes (panel B) are measured in percentages (0-100). 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).

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

	Comparison	GCP	Estimated Effect
	(1)	(2)	(3)
A. Personal Outcomes			
Outside Israel	2.91	6.18	3.42**
			(1.30)
Marriage	36.18	37.52	0.93
			(2.84)
Has Kids	21.45	19.09	-1.86
			(2.25)
B. Partner Outcomes			
GCP Participant	1.11	9.09	8.28***
			(2.31)
UPET Total Score	586.49	609.78	28.89***
			(10.15)
Same Class	5.00	5.56	0.35
			(2.32)
Same GCP Class	0.00	5.56	5.38***
			(1.63)
Students	550	550	1,100
Schools	138	11	149

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

Notes: The table shows estimates for the impact of GCP participation on personal outcomes. The sample includes only students who participated in 8^{th}-grade Metzav tests, about half of the students in cohorts of high-school graduates in 2006-2010. All outcomes, except for the partner's UPET score, are measured in percentages (0-100). The partner's UPET score is measured on a scale of 200-800, and its effect is estimated using a sample of married individuals. 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).

Table 8: Heterogeneity of The Impact of GCP

	Female	High SES	High ability	Since midschool
	(1)	(2)	(3)	(4)
A. Matriculation				
Mean Composite Score	0.78	0.09	0.06	-0.54
	(0.69)	(0.62)	(0.71)	(0.51)
B. Post-secondary Education				
BA Attainment	-2.93	7.07*	-0.93	-1.77
	(4.60)	(4.28)	(4.62)	(3.47)
BA, Double Major	-3.02	1.34	3.70	2.71
	(5.01)	(4.37)	(4.52)	(4.15)
C. Labor-Market, 2020				
Annual Earnings	-0.03	-0.05	0.24	0.27
	(0.20)	(0.18)	(0.18)	(0.18)
Knowledge Economy	-8.56	0.28	3.23	5.73
	(5.79)	(5.03)	(5.28)	(4.64)
D. Personal, 2020				
Out of Israel	-2.01	2.43	2.91	1.45
	(2.52)	(2.67)	(2.55)	(2.31)

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

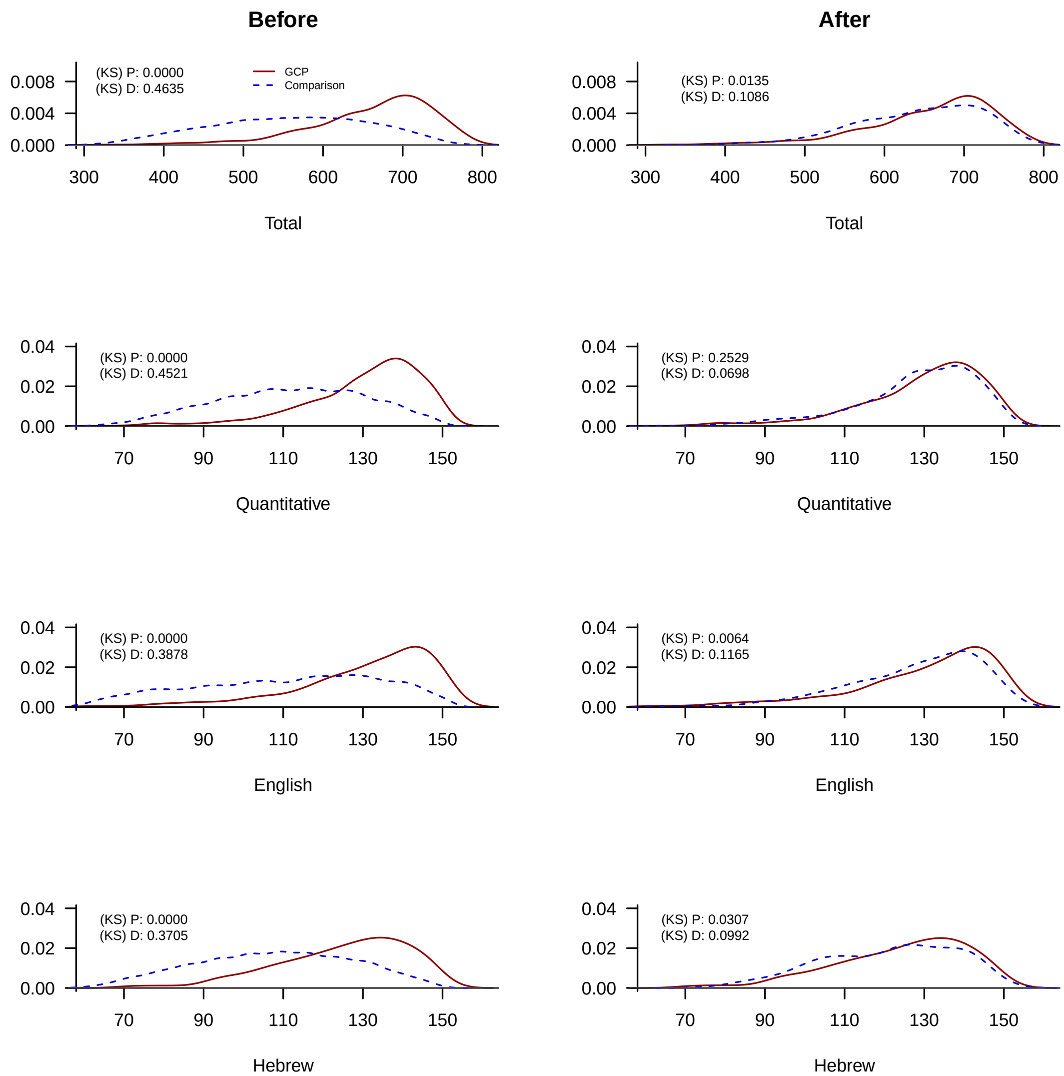
Notes: The table shows estimates for the heterogeneity of the impact of GCP participation on outcomes. The sample includes only students who participated in 8^{th}-grade Metzav tests, about half of the students in cohorts of high-school graduates in 2006-2010. Columns (1)-(4) show the estimated difference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

Online Appendices

Appendix A Data Appendix

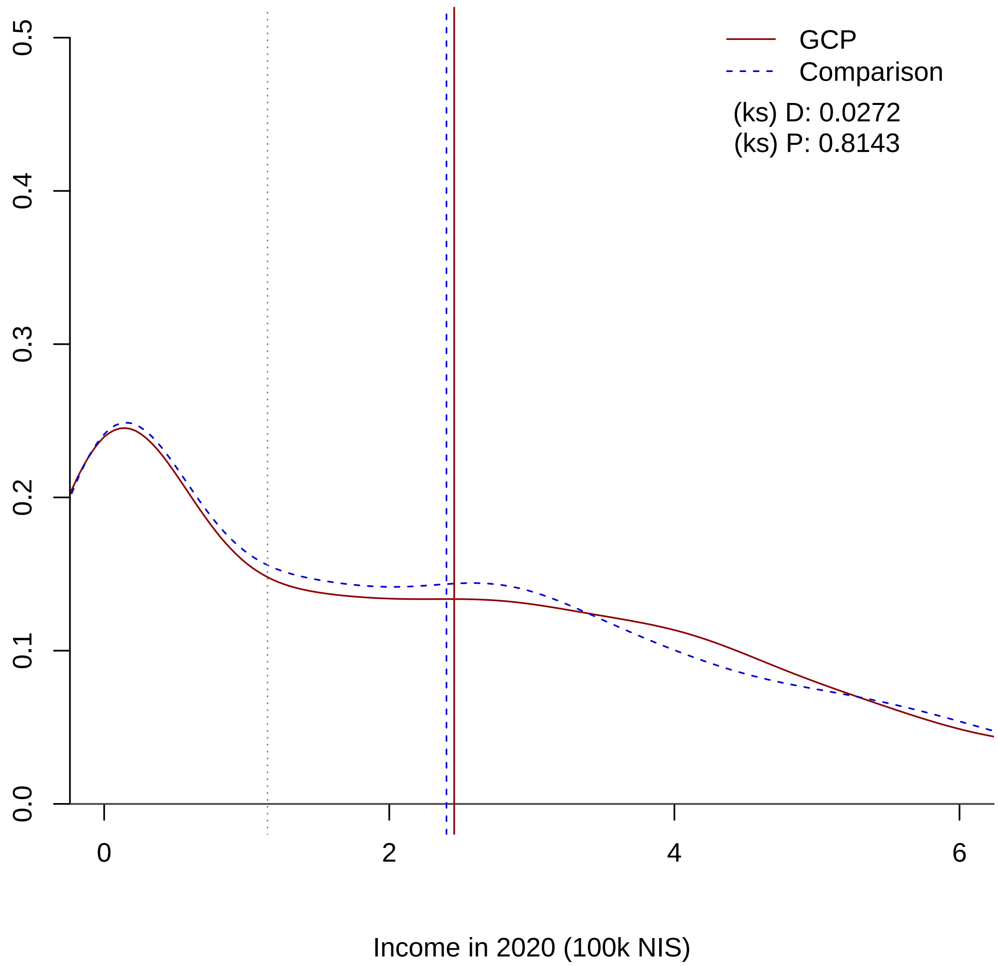
We use several panel datasets from Israel's Central Bureau of Statics (CBS). CBS allows restricted access to this data in their protected research lab. The underlying data sources include the following. The population registry data consists of a fictitious individual national I.D. number that appears 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 10^{th}-12^{th} grades. Another source is 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.

Appendix Figure A1: UPET Scores, Before and After Matching



Notes: This figure plots the UPET scores distribution by groups. The solid red line represents the sample of GCP students, and the blue dashed line represents the matched comparison group (which includes non-GCP students from other cities). The graphs on the left show the distributions before the matching, and those on the right show the distributions after the matching. The graphs also show the Kolmogorov–Smirnov test for the equality of the probability distributions. The sample includes students who participated in the 8^{th}-grade Metzav 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: 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 graphs also show the Kolmogorov–Smirnov test for the equality of the probability distributions. The dotted black line represents the comparison group pull, which includes non-gifted students. The sample includes students who participated in the UPET (at any age) from the cohorts of high-school graduates in 1992-2005.

Appendix Table A1: The Impact of GCP on Matriculation Test Scores

	Comparison	GCP	Estimated Effect
	(1)	(2)	(3)
Mean Composite Score	87.70	87.88	-0.04
			(0.35)
Math	86.75	83.90	-3.48***
			(0.92)
Hebrew	84.56	86.31	1.13**
			(0.47)
English	89.58	91.58	1.39**
			(0.51)
Bible	85.86	86.10	-0.59
			(0.54)
Literature	79.02	81.27	1.45**
			(0.72)
Citizenship	83.24	83.69	-0.32
			(0.64)
History	83.33	84.06	0.14
			(0.68)
Students	550	550	1,100
Schools	138	11	149

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

Notes: The table shows estimates for the impact of GCP participation on matriculation (Bagrut) test scores. The sample includes only students who participated in 8^{th}-grade Metzav 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 A2: Comparison of GCPs to Comparison Students' Classes

	Comparison	GCP	Estimated Effect
	(1)	(2)	(3)
Number of students	34.02	27.31	-6.70***
			(0.30)
5 Credits in Physics (%)	24.60	50.66	26.06***
			(1.07)
UPET Score	586.54	656.71	70.17***
			(2.50)
Father Education	14.22	15.55	1.32***
			(0.07)
Mother Education	14.33	15.28	0.95***
			(0.06)
BA (%)	64.94	77.39	12.45***
			(0.71)
Students	550	550	1,100
Schools	138	11	149

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

Notes: The table shows the difference between GCPs and the comparison group's students' classes. The sample includes only students who participated in 8^{th}-grade Metzav 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).

Appendix Table A3: Comparison of Localities with and without GCPs

	Comparison	GCP	Estimated Effect
	(1)	(2)	(3)
Number of students	810.25	3306.81	2496.57***
			(123.50)
5 Credits in Physics (%)	9.27	8.51	-0.77***
			(0.18)
UPET Score	562.09	560.50	-1.59
			(1.60)
Father Education	13.45	13.52	0.07
			(0.05)
Mother Education	13.54	13.36	-0.18***
			(0.04)
BA (%)	46.75	44.63	-2.12***
			(0.60)
Students	550	550	1,100
Schools	138	11	149

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

Notes: The table shows the difference between localities with and without GCPs. The sample includes only students who participated in 8^{th}-grade Metzav 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 A4: The Impact of GCP on Labor-market Outcomes in 2020, Among Individuals who Live in Israel

	Comparison	GCP	Estimated Effect
	(1)	(2)	(3)
A. Annual Earnings			
Earnings (100K NIS)	1.70	1.69	0.01
			(0.11)
Log Earnings	11.60	11.65	0.06
			(0.09)
Earnings Rank	59.57	60.13	0.83
			(1.76)
Top 10% Earners	25.84	25.19	0.33
			(2.57)
Top 1% Earners	5.99	5.23	-0.59
			(1.41)
B. Employment			
Salaried Employment	84.83	82.75	-2.10
			(2.29)
Self Employment	6.93	4.84	-1.88
			(1.43)
Knowledge Economy	42.13	41.86	-0.38
			(2.97)
Tech Services	32.40	33.33	0.85
			(2.82)
Tech Manufacturing	2.25	1.94	-0.13
			(0.94)
Academic	7.49	6.59	-1.10
			(1.66)
Students			1,050
Schools			149

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

Notes: The table shows estimates for the impact of GCP participation on labor-market outcomes. The sample includes only students who participated in 8^{th}-grade Metzav tests, about half of the students in cohorts of high-school graduates in 2006-2010. Earnings (panel A) are measured in 100K NIS. Employment outcomes (panel B) are measured in percentages (0-100). 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: Robustness of the Results, Alternative Matching Specifications

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
A. Matriculation							
Mean Composite Score	-0.04	0.45	0.35	-0.17	0.49	0.14	0.12
	(0.35)	(0.39)	(0.36)	(0.35)	(0.36)	(0.36)	(0.40)
B. Post-secondary Education							
BA Attainment	1.39	5.67**	4.84*	2.26	2.56	3.66	-2.38
	(2.37)	(2.38)	(2.43)	(2.37)	(2.42)	(2.39)	(2.50)
BA, Double Major	6.82***	9.59***	7.69***	6.92***	8.70***	6.92***	8.85***
	(2.40)	(2.33)	(2.46)	(2.41)	(2.44)	(2.45)	(2.51)
C. Labor-market							
Earnings (100K NIS)	-0.06	-0.03	0.02	0.05	-0.04	0.06	-0.09
	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.11)
Knowledge Economy	-1.61	0.01	-0.45	0.46	-0.46	0.20	-2.33
	(2.89)	(2.82)	(2.90)	(2.86)	(2.91)	(2.81)	(3.08)
D. Personal							
Out of Israel	3.42**	3.68***	2.84**	2.23	1.84	2.62*	2.25
	(1.30)	(1.19)	(1.40)	(1.37)	(1.37)	(1.33)	(1.42)
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	1,100	1,130	1,054	1,094	1,126	938	960

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

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 pull. The sample includes only students who participated in 8^{th}-grade Metzav 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 A6: Robustness of the Results, Matching on Alternative Sets of Variables

	(1)	(2)	(3)	(4)	(5)	(6)
A. Matriculation						
Mean Composite Score	-0.04	0.89**	0.28	0.51	1.13***	0.92**
	(0.35)	(0.36)	(0.29)	(0.34)	(0.36)	(0.37)
B. Post-secondary Education						
BA Attainment	1.39	7.31**	1.03	-0.81	2.38	3.30
	(2.37)	(2.82)	(2.04)	(2.29)	(2.29)	(2.35)
BA, Double Major	6.82***	3.40	4.07	7.14***	5.80**	7.52***
	(2.40)	(2.09)	(2.56)	(2.41)	(2.45)	(2.40)
C. Labor-market						
Earnings (100K NIS)	-0.06	0.12	-0.10	-0.09	0.12	0.02
	(0.10)	(0.09)	(0.12)	(0.10)	(0.10)	(0.10)
Knowledge Economy	-1.61	3.34	0.73	-1.75	2.09	1.38
	(2.89)	(2.76)	(2.94)	(2.87)	(2.81)	(2.80)
D. Personal						
Out of Israel	3.42**	1.67*	2.47**	3.22**	3.08**	2.47*
	(1.30)	(0.96)	(1.14)	(1.30)	(1.24)	(1.26)
Cohorts	06-10	09-13	06-10	06-10	06-10	06-10
Ability Measure	8^{th}-grade	5^{th}-grade	UPET	8^{th}-grade	8^{th}-grade	8^{th}-grade
Matriculation indicators:						
English & Math	+	+	+	+	+	
Scientific Elective	+	+	+	+		
Other Elective	+	+	+			
Number of observations	1,081	1,058	1,145	1,124	1,135	1,154

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

Notes: The table shows estimates for the impact of GCP participation on outcomes, using different sets of variables in the matching specification: different ability measures (5^{th}-grade Metzav test scores, 8^{th}-grade Metzav test scores and quantitative UPET scores), and different sets of matriculation indicators. The sample includes only students who participated in the relevant tests, from the cohorts mentioned on 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).

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)
Age $<$ 17	-14.22	29.97	-1.48	8.01
	(10.45)	(33.00)	(1.96)	(6.21)
Age $=$ 18	1.92	8.25	-0.53**	0.93
	(1.42)	(5.46)	(0.27)	(1.02)
19 $\geq$ Age $\geq$ 21	1.48	11.26**	-3.94***	-1.36
	(1.04)	(5.36)	(0.19)	(1.01)
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 regressions to predict the UPET score of students. 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 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 parenthesis.

Appendix Table A8: Robustness of the Results, 1992–2005 Graduates

	(1)	(2)	(3)	(4)
A. Education				
BA Attainment	-0.46	-0.28	-0.65	-1.04
	(1.09)	(1.11)	(0.92)	(0.94)
	93.47	93.16	90.52	91.17
BA, Double Major	5.55***	4.14**	4.74***	5.15***
	(1.91)	(1.93)	(1.45)	(1.44)
	24.48	26.62	25.23	25.15
C. Labor-market				
Earnings (100K NIS)	0.22	0.04	0.10	0.00
	(0.15)	(0.20)	(0.11)	(0.10)
	2.48	2.66	2.45	2.50
Knowledge Economy	-1.32	1.61	0.32	-0.54
	(2.07)	(2.09)	(1.53)	(1.53)
	41.34	39.38	37.05	38.43
D. Personal				
Out of Israel	2.00	1.63	2.22**	2.69**
	(1.41)	(1.43)	(1.06)	(1.05)
	11.06	11.39	10.98	10.54
Only quantitative score	+		+	
Only early takers	+	+		
Number of observations	2,206	2,194	3,860	3,872

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

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).

Appendix Table A9: Heterogeneity of The Impact of GCP on Outcome, By Gender

	Males ($N = 668$)		Females ($N = 432$)		Difference
	Baseline	Effect	Baseline	Effect	
	(1)	(2)	(3)	(4)	(5)
A. Matriculation					
Mean Composite Score	88.06	-0.35	87.15	0.43	0.78
		(0.47)		(0.55)	(0.69)
B. Post-secondary Education					
BA Attainment	73.65	2.53	78.70	-0.40	-2.93
		(3.20)		(3.77)	(4.60)
BA, Double Major	15.57	7.99***	18.06	4.97	-3.02
		(3.07)		(3.89)	(5.01)
C. Labor-Market, 2020					
Earnings (100K NIS)	1.99	-0.05	1.13	-0.08	-0.03
		(0.15)		(0.17)	(0.20)
Knowledge Economy	47.90	1.73	30.09	-6.83	-8.56
		(3.82)		(4.58)	(5.79)
D. Personal, 2020					
Out of Israel	2.99	4.20**	2.78	2.20	-2.01
		(1.70)		(2.04)	(2.52)

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

Notes: The table shows estimates for the heterogeneous impact of GCP participation on outcomes by gender. The sample includes only students who participated in 8^{th}-grade Metzav test scores, about half of the students in cohorts of high-school graduates in 2006-2010. Columns (1) and (3) show the baseline means among the matched comparison group for males and females. Columns (2) and (4) show the estimated effect for males and females and their standard error. Column (5) shows the estimated difference between the effects (δ from equation (3)) and its standard error (clustered at the school level).

Appendix Table A10: Heterogeneity of The Impact of GCP on Outcome, By SES

	Low-SES ($N = 445$)		High-SES ($N = 655$)		Difference
	Baseline	Effect	Baseline	Effect	
	(1)	(2)	(3)	(4)	(5)
A. Matriculation					
Mean Composite Score	84.72	-0.09	89.65	-0.01	0.09
		(0.54)		(0.43)	(0.62)
B. Post-secondary Education					
BA Attainment	67.71	-2.80	81.04	4.27	7.07*
		(3.64)		(2.91)	(4.28)
BA, Double Major	13.00	6.02*	18.96	7.36**	1.34
		(3.34)		(3.00)	(4.37)
C. Labor-Market, 2020					
Earnings (100K NIS)	1.60	-0.03	1.69	-0.08	-0.05
		(0.14)		(0.13)	(0.18)
Knowledge Economy	37.67	-1.77	43.12	-1.49	0.28
		(4.02)		(3.54)	(5.03)
D. Personal, 2020					
Out of Israel	2.24	1.98	3.36	4.41***	2.43
		(1.95)		(1.57)	(2.67)

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

Notes: The table shows estimates for the heterogeneous impact of GCP participation on outcomes by socio-economic status (SES). Low (high) SES includes students whose father's years of education are less than 15 (15 or more). The sample includes only students who participated in 8^{th}-grade Metzav test scores, about half of the students in cohorts of high-school graduates in 2006-2010. Columns (1) and (3) show the baseline means among the matched comparison group for low and high SES students. Columns (2) and (4) show the estimated effect for males and females and their standard error. Column (5) shows the estimated difference between the effects (δ from equation (3)) and its standard error (clustered at the school level).

Appendix Table A11: Heterogeneity of The Impact of GCP, By Giftendess

	Lower ($N = 335$)		Higher ($N = 765$)		Difference
	Baseline	Effect	Baseline	Effect	
	(1)	(2)	(3)	(4)	(5)
A. Matriculation					
Mean Composite Score	82.08	-0.08	90.41	-0.02	0.06
		(0.68)		(0.39)	(0.71)
B. Post-secondary Education					
BA Attainment	62.84	2.05	82.02	1.12	-0.93
		(4.20)		(2.67)	(4.62)
BA, Double Major	14.21	4.18	17.71	7.88***	3.70
		(3.80)		(2.75)	(4.52)
C. Labor-Market, 2020					
Earnings (100K NIS)	1.28	-0.23	1.83	0.01	0.24
		(0.14)		(0.12)	(0.18)
Knowledge Economy	28.42	-3.91	47.14	-0.67	3.23
		(4.58)		(3.25)	(5.28)
D. Personal, 2020					
Out of Israel	2.73	1.35	3.00	4.26***	2.91
		(2.08)		(1.44)	(2.55)

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

Notes: The table shows estimates for the heterogeneous impact of GCP participation on outcomes by level of giftedness. Low (high) giftedness group includes students whose sum of standardized Metzav scores (in all four subjects) equals four standard deviations or more. The sample includes only students who participated in 8^{th}-grade Metzav test scores, about half of the students in cohorts of high-school graduates in 2006-2010. Columns (1) and (3) show the baseline means among the matched comparison group for low and high giftedness students. Columns (2) and (4) show the estimated effect for males and females and their standard error. Column (5) shows the estimated difference between the effects (δ from equation (3)) and its standard error (clustered at the school level).

Appendix Table A12: Heterogeneity of The Impact of GCP, By Length of Participation

	Only high ($N = 919$)		Since middle ($N = 181$)		Difference
	Baseline	Effect	Baseline	Effect	
	(1)	(2)	(3)	(4)	(5)
A. Matriculation					
Mean Composite Score	87.70	0.14	NA	-0.40	-0.54
		(0.40)		(0.50)	(0.51)
B. Post-secondary Education					
BA Attainment	75.64	1.98	NA	0.21	-1.77
		(2.67)		(3.39)	(3.47)
BA, Double Major	16.55	5.91**	NA	8.63**	2.71
		(2.69)		(3.50)	(4.15)
C. Labor-Market, 2020					
Earnings (100K NIS)	1.65	-0.15	NA	0.12	0.27
		(0.11)		(0.15)	(0.18)
Knowledge Economy	40.91	-3.52	NA	2.22	5.73
		(3.20)		(4.12)	(4.64)
D. Personal, 2020					
Out of Israel	2.91	2.94**	NA	4.39**	1.45
		(1.46)		(1.83)	(2.31)

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

Notes: The table shows estimates for the heterogeneous impact of GCP participation on outcomes by the length of participation in a GCP (since mid-school relative to participation only in high school). The sample includes only students who participated in 8^{th}-grade Metzav test scores, about half of the students in cohorts of high-school graduates in 2006-2010. Columns (1) and (3) show the baseline means among the matched comparison group for those enrolled in the GCP since high-school and those enrolled since mid-school. Columns (2) and (4) show the estimated effect for males and females and their standard error. Column (5) shows the estimated difference between the effects (δ from equation (3)) and its standard error (clustered at the school level).

Appendix Table A13: Heterogeneity of The Impact of GCP, Using the Sample of 1992–2005 Graduates

	Female	High SES	High ability
	(1)	(2)	(3)
A. Post-secondary Education			
BA Attainment	-0.25	-1.47	3.12
	(2.12)	(1.89)	(2.00)
BA, Double Major	6.64*	0.63	-6.26*
	(3.88)	(3.40)	(3.44)
B. Labor-Market, 2020			
Annual Earnings	-0.31	0.20	0.27
	(0.27)	(0.24)	(0.18)
Knowledge Economy	-1.43	-1.83	1.50
	(4.17)	(3.67)	(3.67)
C. Personal, 2020			
Out of Israel	3.01	3.19	-2.26
	(2.83)	(2.40)	(2.52)

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

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 1992–2005. 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 from equation (3)) and its standard error (clustered at the school level).