Bank of Israel **Research Department**

Evaluating the Effectiveness of the "Regular Bagrut (Matriculation) Track" (Mabar) Program

Sefi Bahar*

Discussion Paper 2020.10
September 2020

Bank of Israel– http://www.boi.org.il

* Sefi Bahar, Research Department – Sefi.Bahar@boi.org.il

Thanks to Yotam Nir, Noam Zussman, Yoav Friedmann, Shai Tsur, Adi Brender and Adi Shani for their helpful comments and advice in the course of this work; to staff of the Research Information Accessibility Team of the Department of International Relations and Statistical Coordination at the Central Bureau of Statistics—for preparing the data and making it accessible in the research room.

חטיבת המחקר, בנק ישראל ת"ד 780 ירושלים 91007
Research Department, Bank of Israel. POB 780, 91007 Jerusalem, Israel

הערכת האפקטיביות של התוכנית "מסלול בגרותי רגיל" (מב"ר)

ספי בכר

תקציר

המחקר בחן את השפעתה של התוכנית "מסלול בגרותי רגיל" (מב"ר) בשנים 2010–2015 על משתני תוצאה בטווח הקצר. מב"ר היא תוכנית תלת-שנתית, המופעלת בבתי ספר תיכוניים במטרה לצמצם את הנשירה ולהעלות את שיעורי הזכאות לתעודת בגרות בקרב תלמידים מרקע חברתי-כלכלי חלש, שהישגיהם הלימודיים בחטיבת הביניים נמוכים. כדי להתגבר על בעיית הסלקציה בכניסה לתוכנית ניצלנו את עודף הביקוש ואת תנאי הקבלה, והשתמשנו בשיטת Propensity Score Matching ליצירת קבוצת השוואה של מועמדים שנתוניהם התאימו לכיתות מב"ר אך הם לא נכללו בהן ולמדו בכיתות רגילות. תוצאות המחקר מלמדות שהשתתפות בכיתת מב"ר מביאה לירידה של 3.2 נקודות אחוז בשיעור הנשירה (כ-40%) ולעלייה של 4.8 נקודות אחוז בשיעור הזכאות לתעודת בגרות. ההשפעות חזקות יותר בקרב בנים ובקרב תלמידי החינוך הערבי: אצלם התוכנית מביאה להפחתה של 6.2 נקודות אחוז בנשירה (70%) ולשיפור של 9 נקודות אחוז בזכאות לתעודת בגרות (20%). לא נמצאה השפעה של התוכנית על תלמידי זרם החינוך הממלכתי-דתי.

Evaluating the Effectiveness of the "Regular Bagrut (Matriculation) Track" ("Mabar") Program

Sefi Bahar

Abstract

The research examined the impact of the "Regular Bagrut (matriculation) Track" ("Mabar") in 2010–15 on result variables in the short term. Mabar is a 3-year program operated in high schools to reduce dropout rates and to increase the eligibility rates for the Bagrut matriculation diploma among students from a weaker socioeconomic background, whose schooling achievements in middle school were low. In order to overcome the selection problem in the entry to the program, we utilized the excess demand and acceptance criteria, and used Propensity Score Matching to create a control group of candidates whose data matched Mabar classes but were not included in those classes, rather studying in regular classes. The results of the research indicate that participation in a Mabar class leads to a decline of 3.2 percentage points in the dropout rate (approximately 40 percent) and to an increase of 4.8 percentage points in the eligibility rates for the Bagrut matriculation diploma. The effects are stronger among boys and among students in the Arab school system, where the program leads to a decline of 6.2 percentage points in the dropout rate (70 percent) and to an improvement of 9 percentage points in the eligibility rates for the Bagrut matriculation diploma (20 percent). No effect of the program was found on students in the State Religious school system.

1. Background and literature overview

This study involves the regular matriculation track (hereinafter: "Mabar" (the Hebrew acronym for "Regular Bagrut (matriculation) Track"))—a remedial education program with the primary goals of reducing dropout rates and increasing eligibility rates for the Bagrut matriculation diploma, as dropping out of school negatively impact both individuals and society as a whole.[1]

Mabar is a three-year program for high school students from a low socioeconomic background, with low scholastic achievement in earlier education stages. The program has been running for over two decades, since the mid-1990s, and in 2017 some 6 percent of high school students took part in this program.[2]

The importance of this program stems from its being designed for students from a low socioeconomic background, which has been found to be correlated with low scholastic achievement (Jensen, 2013). Israel has one of the highest socioeconomic achievement gaps in the world (OECD, 2019) and this gap grows with each higher education stage. When reaching high school, students from a low socioeconomic background are at high risk of dropping out of school, or completing their course of study without a diploma (Bowers et al., 2012; Dovrin, 2015).

Despite the significant importance of the Mabar program, by virtue of its objective of reducing gaps and inequality in society, the large number of participants, and its high annual cost (NIS 350 million), to date no study has been conducted of the program's effectiveness in achieving its goals. The key reasons for this are methodology issues, reflected in biased selection of program participants, and the many elements applied in the intervention, which make it challenging for researchers to isolate the contribution of each element to any improvement in achievement. In this study, we applied advanced research methods to reduce the bias due to non-random selection of program participants, by comparing their achievement to that of students and graduates with similar features who did not take part in the program. This is intended to verify whether the program achieves

[1] There is extensive literature regarding the negative implications of school dropout, including: Higher likelihood of idleness, low earning capacity, impaired health condition, higher risk of incarceration and so forth. For more details see Belfield & Levin, 2007; Chapman, Laird, Ifill, & Kewal Ramani, 2011; Lleras-Muney, 2005.

[2] Calculated across all high school students; Mabar students account for 9 percent of all students in schools where the program may operate (regular, non-vocational schools where matriculation exams are offered).

its objectives, and to help the Ministry of Education in setting policy in future, designed to reduce the school dropout rate and to increase the matriculation eligibility rate.

Many studies involve remedial education, addressing learning issues. These studies focus on the impact of such programs on scholastic achievement, on the effectiveness of various types of intervention (additional hours, exercises using computer software, monetary incentive to students and so forth) and their impact on the target population in various educational stages (elementary school, middle school, high school and so forth).[3]

In Israel, two studies have studied the impact of a remedial education program for students with low scholastic achievement, named "Bagrut 2000", which included additional teaching hours, smaller class size and study in smaller groups. Lavy & Schlosser (2005) found that the probability of students in this program to achieve matriculation increased by 13 percent and that schools that participated in this program improved their matriculation eligibility rate by 3.3 percentage points. Another study (Lavy, et al. 2018) studied the long-term impact of this program—when students were 33 years old. The study outcome showed that program participants have completed more years of education, have increased their annual salary by 4 percentage points and have increased the number of months worked per year by 1.5 months on average. The program was not found to have an impact on the likelihood of admission to higher education.

The Mabar program is fundamentally different from all the programs covered in literature, both in terms of duration of intervention (3 years) and in terms of number of elements applied in this intervention. Mabar is a remedial teaching program classified as "Alternative educational program", i.e., intervention conducted in a regular school, by a class offering an alternative educational program (Hefner-Packer, 1991; Freeman & Simonsen, 2015). All participants in the program benefit from studying in a small class, from additional teaching hours and from a team that has received specific training. The different elements applied in the Mabar program were found by many studies to have an impact on scholastic achievement, in particular for students from a weak socioeconomic background. The impact of the change in class size on student achievement is one of the questions that has been facing researchers in recent decades, and early studies of this subject have found a positive impact—in particular for students from a weak background.

[3] For more information see: (Banerjee et al., 2010; Banerjee et al., 2016; Bessho et al., 2019; Eran et al., 2017; Banerjee et al,. 2007).

Thus, for example, the STAR project included a field experiment that involved random assignment of students to classes of different sizes, looking at the long-term impact of the difference in class size on student achievement. Indeed, the class size was found to have an impact, especially on achievement of African-American students. In Israel, Angrist and Lavy found that reducing the number of students per class resulted in improved achievement in reading and mathematics among 4th and 5th grade students, and in particular among students from a weak socioeconomic background (Angrist & Lavy, 1999). In contrast, Kingdon and Altinok found, when studying the impact of class size on student achievement across 47 countries, that only 14 of the countries showed a significant negative correlation between class size and student achievement (Altinok & Kingdom, 2012).

The impact of increase in number of study hours is the single most studied topic for remedial education programs, and the literature review above shows that such increase has a positive impact on student achievement, in particular for students from a weak socioeconomic background. The third element included in the intervention in this program is application of different pedagogic methods, in particular the self-efficacy approach, which was found to have a positive impact on student achievement. A long-term study of students aged 13–19 showed that early exposure to the self-efficacy pedagogic approach resulted in significant improvement in achievement, in particular for students from a weak socioeconomic background (Caprara, et al. 2011).

2. Description of the "Regular Bagrut (matriculation) Track" (Mabar) program

The Ministry of Education has been working diligently to reduce the school dropout rate in the school system across different education levels, and to reduce the scholastic gaps that adversely impact students from a low socioeconomic background. This is done through various programs designed to promote perseverance and to make up any study gaps accumulated by students. One of the Ministry organs used to realize its targets with regard to reducing the dropout rate and to reducing gaps is Division for Education of Children and Youth at Risk, which strives to "maximize individual potential and formulate the future perception of children and youth at risk".[4] This division operates the Mabar

[4] https://edu.gov.il/minhalpedagogy/minhal/units/Pages/risk.aspx

program in Shachar (Educational and Welfare Service) classes, where programs are applied in different education levels designed to "enhance the scholastic achievement of students and increase the number of those eligible for a matriculation diploma among non-affluent populations".[5]

The Mabar program has been in operation since the mid-1990s and is designed to empower students at the emotional, moral and cognitive levels, so as to enable them to achieve matriculation at the optimal quality level, which would allow them to gain admission to higher education (REMA, 2019). This is a three-year program (10^{th}–12^{th} grades) for students from a weak socioeconomic background with low achievement in middle school, and has the following key objectives:

a) Prevent dropouts and increase perseverance of Mabar class students in 10^{th}-12^{th} grades;

b) Significantly increase the matriculation eligibility rate; Improve the matriculation diploma and matriculation eligibility;

c) Improve the matriculation diploma, from 3 curriculum units to 4 and 5 curriculum units in English and in mathematics;

d) Develop values of belonging and social engagement of students in the community;

e) Develop tools to increase motivation, self image, scholastic efficacy and emotional coping.

In this program, students receive multiple interventions, including:

- Reduced number of students per class (to 18–25);
- Increased number of teaching hours per student, compared to students in regular classes (by 40 percent);
- Custom training delivered to teaching staff (using the self-efficacy approach for students and teachers).

[5] Ibid. In addition to Mabar, during the study period Shachar classes also operated the ETGAR program, designated for students with lower scholastic achievement, with the key goals of preventing dropouts and achieving at least 16 curriculum units for matriculation. This program is not part of this study, since it is not subject to the assumed diversion of excess demand to regular classes, as elaborated in the section on student selection for this program.

For a school to open a Mabar class, the school must meet several pre-conditions.[6] The school must be recognized as offering matriculation; offering diverse study subjects; maintaining proper examination conduct (fewer than 2 percent disqualified tests); maintaining a pedagogic process resulting in improved student perseverance. Preference is given to schools with over 50 percent of students in nurture index deciles 5-10.[7]

Preconditions for students are:

- **Low scholastic achievement in 9th grade** – Low achievement is defined as an average grade of 50 to 60 in four major subjects (mathematics, English, Hebrew language and history) and no more than 4 Fail grades.
- **Socioeconomic background** – Preference is given to students in cultivation decile 5 or higher of the benchmark calculated by the Division. Criteria that impact the cultivation benchmark grade are: Years of study of father and mother, number of siblings, marital status of parents, employment status of parents, residence in areas of national priority and up to 10 years living in Israel.

The annual program cost per student is estimated at an additional NIS 13,000 over the standard budget per high school student. In 2018, the program included 26,000 students in 1,135 classes across 434 schools. The annual incremental cost of this program is estimated at NIS 350 million.

3. Database and description of study population

A. Database

This study was conducted at the Central Bureau of Statistics study room, using anonymous data for individuals and schools[8], based on many files of administrative data used to create the integrated study database. Below we present the data files, groups of students excluded from the study and the reasons for such exclusion.

---

[6] Source: "Procedures and Preparation for Opening Mabar Classes"—booklet on Ministry of Education website.

[7] A cultivation decile is a socioeconomic benchmark calculated by the Ministry of Education. The higher the value is on this benchmark, the weaker the socioeconomic standing of the student, hence the higher their need for cultivation.

[8] Data files did not include any identifying data, i.e., no reference to student ID, school ID or any other identifying element. Therefore, and for the sake of clarity, any reference to identifying data below is reference to fictitious identifying data (coded by the Central Bureau of Statistics).

The study included all high school students who started 10th grade in 2010–15 in the official State education school system or supervised thereby, in the State Religious education school system, in the Jewish and in the Arab sectors. Students in the unofficial school system, in special education and in Jewish ultra-Orthodox education were excluded, since in most years of the study period the Mabar program was not offered at such schools, and due to their unique attributes. We also excluded Druze, Circasian, Bedouin and Christian schools—due to the small number of such schools where the program was offered. (These constraints were mandated by the Central Bureau of Statistics Confidentiality Committee).

- **"Students – Ministry of Education" file** – This file includes data for all students in various education stages for a given year. The information in this file is at the individual student level, and was used to create the baseline for the integrated database. Fields used for this purpose are: Individual ID, school ID, school year, gender, ethnicity, school grade, corresponding grade, track, matriculation graduation year, parents' education, number of siblings and parents' ID (fictitious).
- **"Actual classes" file** – This administrative file includes information about classes, allowing for identification of student participation in the program (by a combination of school grade, corresponding grade and class type). Identification for the purpose of the program was based on the student's class type when in 10th grade and, if the student did not drop out, he or she must have remained in the same class type across all subsequent high school grades.[9] This means that a student who did not drop out of school received the full benefit offered by the program. This file is continuously available from the Central Bureau of Statistics from 2010 to date. Fields used to generate variables are: class type, school type and school cultivation benchmark grade.
- **"Population Registry – Ministry of Interior" file** – This file includes information about background attributes of the parents and was used to verify data for individual attributes from the Students file. Data items obtained from this file are: parents' education, number of siblings and parents' year of immigration to Israel.

[9] Moving between class types is rare, hence students who moved between class types were excluded from this study.

- **"Meitzav test scores in 8th grade" file** – This administrative file contains information about Meitzav test scores of 8th grade students for 2002–17. The fields used to create the database are: test year in 5th or 8th grade and Meitzav test score for each subject. For each test subject in a given year, student achievement was standardized by percentile, to allow for comparison over time.
- **"Matriculation scores" file** – This administrative file contains information about student achievement on matriculation tests. The data fields obtained from this file are: matriculation eligibility and matriculation scores for various subjects.
- **"Schools" file** – This file includes information about school attributes. The major contribution of this file was to isolation of the study population and to control over macro attributes of schools. The fields used in this study are: school system, legal status (official, recognized, recognized semi-private, etc.), sector, region, reporting entity, matriculation track and school type.
- **"Teaching staff in high schools" file** – This administrative file includes information about attributes of teaching staff in high schools. Fields used in this study are: teaching seniority and full-time position equivalence.

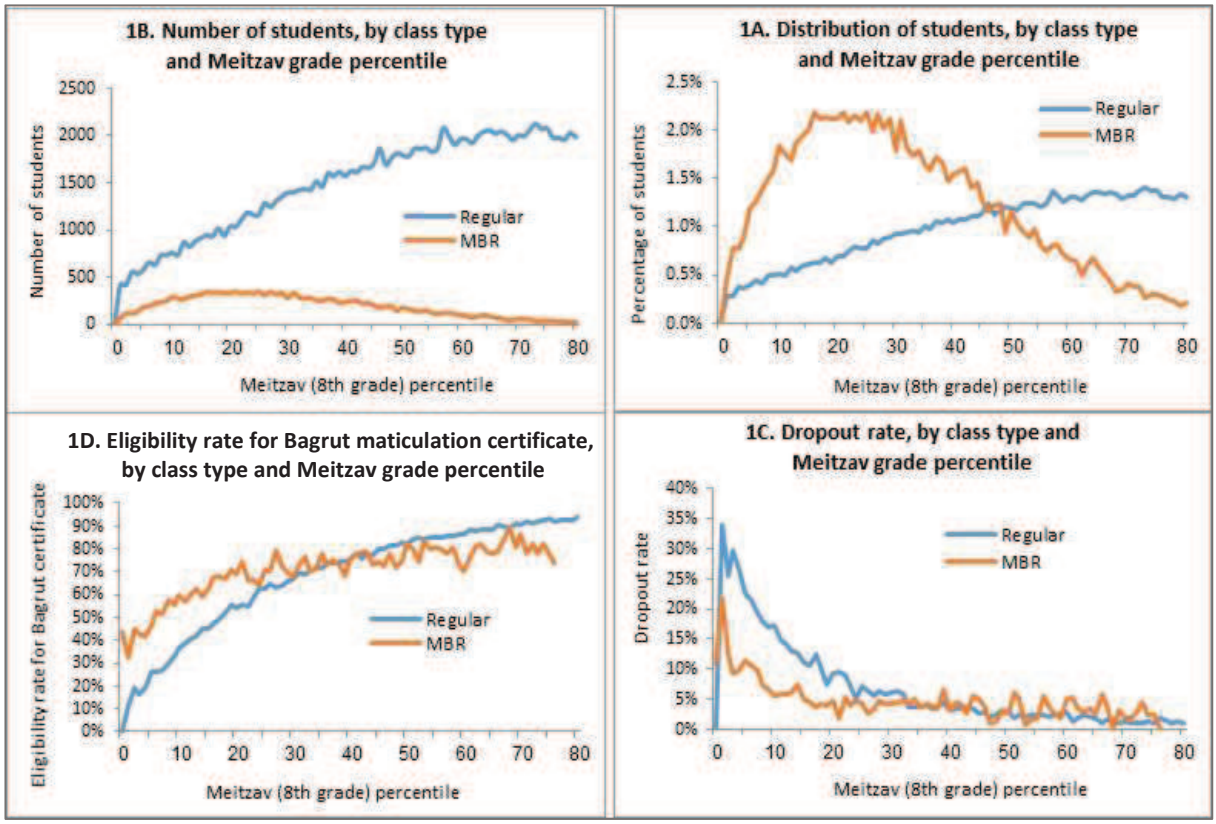
3.1. *Descriptive statistics of study population*

The set of Figures 1 and Table 1 show the statistical attributes of students in Mabar classes and in ordinary classes. The figures show student composition by class type and Meitzav score percentile (X axis of all figures). Figure 1A shows that Mabar students are in lower Meitzav score percentiles than students in ordinary classes: The average 8th grade Meitzav score percentile is 32, compared to 56 for ordinary classes. As shown in Figure 1B, even though Mabar classes have a higher share of students with low Meitzav scores, each Meitzav score percentile has more students in ordinary classes than in Mabar classes. This data will allow us later on to create the peer group for Mabar students, using students with similar attributes in ordinary classes.

Figures 1C and 1D show the dropout rate and the matriculation eligibility rate by Meitzav score percentile and by class type. In this study, the definition of dropout differs from the official definition of the Ministry of Education. A dropout student is any student who failed to continue their studies in the recognized, nonvocational school system. However, the definition of matriculation eligibility is identical to the one used and calculated by the

Ministry of Education.[10] The figures show that students in lower Meitzav score percentiles tend to have a higher dropout rate, but looking at the figure by class type shows that for students with low achievement in 8th grade (30 percent of students with lowest Meitzav scores), the success rate in Mabar classes is higher: They have a lower dropout rate and a higher matriculation eligibility than their peers (in terms of Meitzav scores) assigned to ordinary classes.

Figure 1: Various parameters by class type and percentile composition of Meitzav Test, 8th grade



Review of student background attributes in the various classes shows that the average achievement of Mabar students in the 8th grade Meitzav test is lower by almost one full standard deviation than that of students in ordinary classes. Mabar students come from families with relatively low education level of their parents: The average number of school years for the mother is 12, compared to 13.2 for ordinary classes (a difference of one half of a standard deviation). The description of variance in program participants' attributes

[10] Due to this difference, in this study we define a student as a dropout if they transfer to a lower vocational school system (MIFTAN), welfare facilities (HILA), Yeshiva and so forth—whereas the Ministry of Education would not consider this student to be a dropout, but rather a non-persevering student.

compared to their peer group matches the findings in literature. For example: A study of achievement by students in over 200 US schools where various programs were implemented, found that the intervention population significantly differs from the general population, with a higher percentage of male students, students from minority groups (African-American and Hispanic) and students from single-parent families (Quinn & Poirier, 2006).

As for the difference in school attributes: Table 1 shows that Mabar students come from schools with a relatively high nurture index grade, by about one quarter of a standard deviation compared to schools that do not participate in this program. Allocation of Mabar classes to Arabic speakers is low: They only account for 15 percent of all students in Mabar classes, compared to 27 percent of students in ordinary classes. Allocation to students in State–Religious schools is high: 12 percent of Mabar students, compared to 14 percent of students in ordinary classes. As for the dropout rate, we can see that it is higher for Mabar

Table 1: Various statistical features of students in Mabar classes and in ordinary classes

Variable	Entire sample	Regular classes	Mabar classes
Meitzav test, 8th grade	54.11 (25.674)	56.39 (25.182)	32.00 (19.031)
Years of education – mother	13.10 (3.092)	13.22 (3.122)	12.03 (2.565)
Number of siblings	2.79 (1.644)	2.78 (1.644)	2.87 (1.643)
Percentage of boys	47% (0.499)	47% (0.499)	45% (0.498)
School nurture index grade	5.16 (2.425)	5.10 (2.447)	5.68 (2.149)
Average teacher seniority at the school	19.24 (4.093)	19.19 (4.122)	19.75 (3.785)
Percentage of Arabic speakers	26% (0.440)	27% (0.446)	15% (0.358)
Percentage of students in State-Religious schools	15% (0.355)	14% (0.348)	22% (0.412)
Percentage of those living in outlying areas	33% (0.470)	33% (0.469)	36% (0.480)
Dropout rate	5.41% (0.226)	5.39% (0.226)	5.67% (0.231)
Matriculation eligibility rate	76.3% (0.426)	77.1% (0.420)	68.2% (0.466)
Average grade of tested students on matriculation diploma	84.7 (15.132)	86.0 (14.954)	73.4 (11.609)

students than for students in ordinary classes (by 0.3 percentage point); their matriculation eligibility rate is lower by 9 percentage points compared to ordinary classes, and their average matriculation scores are lower by 12.5 points than for students in ordinary classes.

4. Methodology and identification method

In this study, we wish to test whether attending Mabar classes results in improvement in a host of outcome variables. This calls for estimating the following equation:

$$(1) \quad Y_{ist} = \alpha + \rho' X_{ist} + \gamma \mathrm{T}_i + \delta_t + \varepsilon_i$$

where:

Y_{ist} – *Outcome variable of student i in school s in year t. In this study we tested the dropout rate, matriculation eligibility rate and average matriculation score of those tested in 12th grade.*

X_{ist} – *Set of individual student and school attributes when student was in 10th grade*

T_i – *Dummy variable: 1 if student attended a Mabar class, 0 otherwise.*

δ_t– Set of dummy variables for year and geographic region where the school is located (as categorized by the Ministry of Education), including interaction variables among these variables. The objective of the interaction variable is to isolate changes to program allocation in any particular year in a particular region (such as more or less strict conditions for taking part in the program).

Estimation of equation 1 using the OLS method would result in biased estimates, due to two key issues: (a) Absence of random assignment of students to the intervention program; and (b) High correlation of variables that affect the likelihood of being included in the intervention group and the result variable: Socioeconomic attributes affect both the likelihood of admission to the program and the student's achievement, thereby generating correlation between attending the program and the original disparity (ε_i), which may result in biased estimators. Without proper handling, the selection bias for admission to the program could result in bias in assessing the impact of attending Mabar classes on the various result variables.

The selection bias occurs in two key aspects: First is the student selection for Mabar classes, and second is the selection of schools where such classes are offered. The program is intended for students with attributes correlated with dropping out and low achievement. Therefore, students not in this program are expected to have different attributes. Such evidence was cited in the previous section, describing the study population. The significant difference between the intervention group and the peer group poses a challenge to proving the causality and to reliably estimating the program impact.

Another difficulty is posed by selection of the schools where the program is offered: Schools with a higher percentage of students with weak background attributes are given preference. In some studies it was even asserted that stronger schools avoided offering such programs due to concerns about impact to their image. For all of these reasons, estimating equation (1) using the OLS method may result in biased estimation of program impact.

We may use various methods to reduce the selection bias upon student admission to the program and upon selection of the schools where the program is offered, but in absence of random assignment to intervention and control groups we are unable to totally eliminate the selection bias. One method for addressing the selection issue is the matching method, which allows for comparison between different groups, whereby assignment to the program is based on observed variables ("Unconfoundedness Assumption").

The underlying assumption of the matching method is that assignment to the intervention group is based, as noted above, exclusively on observed variables and by comparing two groups with identical attributes, the only difference being the intervention, we obtain a result similar to that of random assignment to the groups. In this method, each observation is assigned a matching score, which reflects the probability of getting intervention, in order to create groups that prior to the intervention were as similar as possible in their various attributes. This study used the Propensity Score Matching (PSM) method, where a matching score is calculated for each individual, based on observed variables. Since there are two selection factors for assignment to the program (individual level and school level), we used the Multilevel PSM method, and the probability of getting intervention was calculated based on attributes of the student and of their school.

This study uses the commonly used definition in the literature for the Average Treatment effect on the Treated, or ATT: ATT in the estimated effect of the program on those for which the program was intended; hence this effect matters to policymakers more than the average treatment effect on the entire population (Heckman et al., 1997). The program effect is identified based on two key assumptions:

$$(1) \quad Unconfoundedness\ (A.1): Y_1, Y_0 \perp T|X,$$

$$(2) \quad Common\ support\ (A.2): 0 < P(T = 1|X) < 1$$

The combination of assumptions A.1 and A.2 is termed "Strong ignorability" (Rosenbaum & Rubin, 1983). The first assumption (A.1) means that in each cell defined by the observed variables (X), treatment is random and the choice of getting the treatment depends exclusively on the observed variables. That is, within the sub-population defined by the values of observed variables (X), allocation is random. The second assumption (A.2) means that values of the observed variables overlap between the treated group and the control group. This means that against observations in the treated group, we can find observations with similar attributes in the control group (Heckman et al., 1997).

Heckman et al. have shown that the selection bias (B) may be decomposed into 3 different components: B = B1 + B2 + B3 (Heckman et al., 1998). B1 refers to bias caused by lack of overlap in the control variable (X) between the treatment group and the control group. The second bias component (B2) is expected due to incorrect weight assigned to an observation within the common support, while the third bias component, termed "True economic selection bias", is due to selection upon admission to the program due to existence of un-observed variables. Under the first assumption, B3 equals zero and the other bias components are eliminated under the second assumption and alignment of the treated group with the control group using the matching method.

Matching is a popular method, with the underlying principle being matching treatment and control groups with the same background attributes (X). Under the first assumption, in each cell of X the assignment to the treatment group or to the control group is random. Therefore, assuming that all variables that impact selection upon receiving the treatment and the result variable are reflected in X values, one can match to each treatment unit one or more control units with the same values of X. Thus, the only difference between the groups is exposure to the treatment, hence any differences between the two groups may be attributed to the actual treatment.

4.1 *Identification method of students designated for the program*

In order to test the effectiveness of the Mabar program, we use the distinct criteria for admission to Mabar classes and the excess demand for such classes in each year. Due to such excess demand (meaning that the number of students eligible, based on the criteria, for inclusion in the Mabar class is higher than the number of places in such class, as allocated at the school), excess students were assigned to ordinary classes, rather than to Mabar classes, based on a decision made by the educational staff (the counselor and the Shachar administrator). We assume that of those students who meet criteria for entering the program, those with the best prospects of completing their studies were assigned to ordinary classes, and weaker students were assigned to Mabar classes. This assumption is based on interviews with Shachar supervisors at the Ministry of Education and on the structure of incentives for school management, designed to maximize achievements. Appendix 8.1 shows a graphical representation of the assignment method and of the ranking of classes to which students may be assigned in 10^{th} grade.

This assignment method negates the Unconfoundedness Assumption, since getting the treatment is affected by an external factor (the educational staff), which may have also been affected by unobserved variables; this is a bias (B3) whose direction would result in underestimation of the program impact.

Another assumption, which apparently does not hold true in this study, is the Stable Unit Treatment Value Assumption (SUTVA), an assumption whereby the potential results of individual *i* are not affected by treatment of other individuals (Rosenbaum & Rubin, 1983). This assumption does not hold true, because removing the weaker students from the class and assigning them to the treatment class should reduce the negative peer effect of those weaker students on the rest of the class. Moreover, because as part of the treatment, school staff receive training on the use of various teaching methods, having an Mabar class may have a positive spillover effect, due to all teaching staff at the school receiving such training.[11] The combination of these effects, should they occur, would be expected to result in underestimation of the program's impact. Therefore, the results presented in this work may be conservative, for the reasons cited above. When we consider the robustness testing

[11] The training is designated for the teaching staff of the Mabar class, but in actual fact, it is attended by many teachers who are not part of the designated staff of the Mabar class.

of the results, we shall attempt to verify what happens when one tries to isolate the program impact from these effects, and we would expect this to result in a stronger program impact.

In the first stage of this study, we used the criteria for participation in the program to estimate the probability of a student being treated; in the second stage, we tested the average treatment effect on those treated. To estimate the probability of being treated, we used the criteria for participation in the program as set forth in Section 2. Below are details of the variables used to estimate this probability:

- **Meitzav test score, 8th grade** – Adjusted variable (proxy) for student achievement in 9th grade. Scores in 9th grade are one of the key variables affecting admission to the program, but since these are intraschool data (and not available to us), we used the Meitzav test score conducted in 8th grade**,** shortly before making the decision on admission to the program. The Meitzav scores for each subject and year were standardized by population percentile, and the average student achievement in both subjects (mother tongue and mathematics) were introduced as an explanatory variable in the model. Achievement before admission to the program have been used as an explanatory variable in many studies, and this variable was found to be strongly correlated with school dropout (Allensworth & Easton, 2007; Balfanz, et al. 2007; Silver, et al. 2008).
- **Gender –** Dummy variable: 1 if student is male, 0 otherwise.
- **Socioeconomic attributes of the student** – Multiple attributes included in the cultivation benchmark calculated at student level, found to be correlated with the probability of the student dropping out of school (Dalton, et al., 2009; Bowers et al., 2012).
    - **Years of education of mother** – This variable was found to be positively correlated with student success;
    - **Number of siblings** – This variable was found to be negatively correlated with student success;
- **School attributes** – Allocation of Mabar classes to schools is not random, but rather affected by various school attributes. Therefore, we included these attributes as well in calculating the probability of being treated. Below are details of variables we used in the estimation:

- **School nurture index grade** – This benchmark is calculated by the Ministry of Education as indicator of the socioeconomic status of school students. The grade is between 1 and 10. Higher grades for this benchmark are typical of schools where students' families are from the weakest population.
- **School teachers' seniority** – This observed variable was found in previous studies to be positively correlate with school teacher quality (Rivkin, et al. 2005).
- **Supervision and sector** – Combination of supervision and sector variables and a distinction among schools from three school systems: official -Jewish state education, official -Arab state education and State Religious education. The baseline group was defined to be the official -Jewish state education system.
- **Ministry of Education region, school year and interaction between the two** – These variables were designed to discount the effects of variance in allocation between regions, of changing factors between years and of region-level factors that affect student assignment to the program and transfer of "excess demand" to ordinary classes.
- **Dummy variable for outlying areas** – If the school is located in the Northern or Southern regions.

After estimating the probability of participation in the program, based on the observed variables, we considered a large number of algorithms to match the groups (treated and un-treated), so as to arrive at a control group that is as similar as possible in attributes to the treated group. In this study we elected to use the Nearest Neighbor algorithm, with the number of neighbors being up to 5 ($K \leq 5$).[12] We also used the Kernel matching algorithm (bandwidth = 0.06) and after matching the groups, we conducted a t test for variance of the result variable.[13] For further robustness testing, we estimated the program impact using the OLS model using the weights (WLS) calculated from the matching algorithm (Abadie & Imbens, 2002).

[12] We tested all options between 1 and 5, and there was no significant difference in results of sample balancing between the different algorithms; however, the higher the number of neighbors, the more conservative the estimates. We also reviewed various matching algorithms as part of the robustness testing.

[13] All estimates of the Propensity Score Matching were made using the Psmatch2 command (Leuven & Sianesi, 2003).

As the main objective of the program is to reduce the dropout rate and increase the matriculation eligibility rate, we focused on these two variables as the outcome variables in this study. We defined a dropout student as one who started studies in 10th grade and did not appear in the Students file when they should have started in 12th grade. The matriculation eligibility variable was calculated based on Ministry of Education definitions, for the winter testing session of the year subsequent to conclusion of studies.[14]

5. Results

5.1 *Estimation of probability of participation in the Mabar program and group balancing tests*

The first stage in implementation of the PSM method is estimation of student probability of attending an Mabar class; in this study, we estimated this probability using a Probit model. The second stage involves matching the groups, based on the matching algorithms. The key results of the probability estimation are presented in Table 2. These results indicate that all estimates obtained were in the expected direction: Adding the years of education of the mother and increase in Meitzav test score in 8th grade reduce the student probability of attending an Mabar class. The dummy variable for boys had a negative value, because the average Meitzav score percentile for boys in an Mabar class is lower than for girls. As for school-level variables, increase in the school cultivation benchmark grade which students attend, increase in teacher teaching seniority or being part of the State-Religious school system increase the student probability of attending an Mabar class. In contrast, being part of the State-Arab school system reduces the student probability of attending a Mabar class.

To test whether the probability estimation succeeded in creating a comparison group that is similar enough, we conducted balance testing of the observed attributes before and after matching. The most common method for testing balance after matching is Standardized Mean Difference (SMD), because in this method, the measurement method does not have an impact, and the number of observations has a low impact on significance of the result; this is opposed to a t test, which is affected by these factors (Zhang, et al. 2019). Another test of the sample balance is further estimation of the probability function after division

[14] Listed as a separate field in Matriculation file.

into treatment and control groups. The lower the explanatory power of the model, the better the group matching. Table 3 shows a comparison of several variables before and after matching, and several tests for the average variance between the groups.

Table 2*:* Estimation results of PROBIT model for estimating the probability of students to be in Mabar class

Results of probability model	
Meitzav test score, 8th grade	-0.0207*** (0.000)
Years of education – mother	-0.0540*** (0.002)
Gender	-0.0602*** (0.011)
Nurture index	0.108*** (0.004)
Median teacher seniority	0.0075*** (0.002)
Dummy – State-Arab	-0.672*** (0.021)
Dummy – State-Religious	0.0671*** (0.016)
Outlying areas	-0.662*** (0.054)
Constant	0.0396 (0.0717)
Dummy for year and Ministry of Education region with interaction	V
Set of dummy variables for number of siblings	V
Pseudo R2	0.187
Number of observations	141,025

Data in Table 3 show that for all variables, except for Percentage of Arab Students, the groups are well-matched and the variables are below the 5 percent threshold, meaning that there is no significant variance between the groups. The column for t tests shows that there is still a significant difference for some of the variables. We believe that these differences are due to the significant variance between school systems (in particular between Hebrew speakers and Arabic speakers), and because allocations for this program are made separately for each school system. Appendix 8.2 shows the balance tests, where groups are segregated by students in Hebrew- and Arabic-speaking schools. The results indicate that the match quality has improved, all variables are below the 5 percent threshold, and no significant difference was found in t tests compared to the overall sample, hence

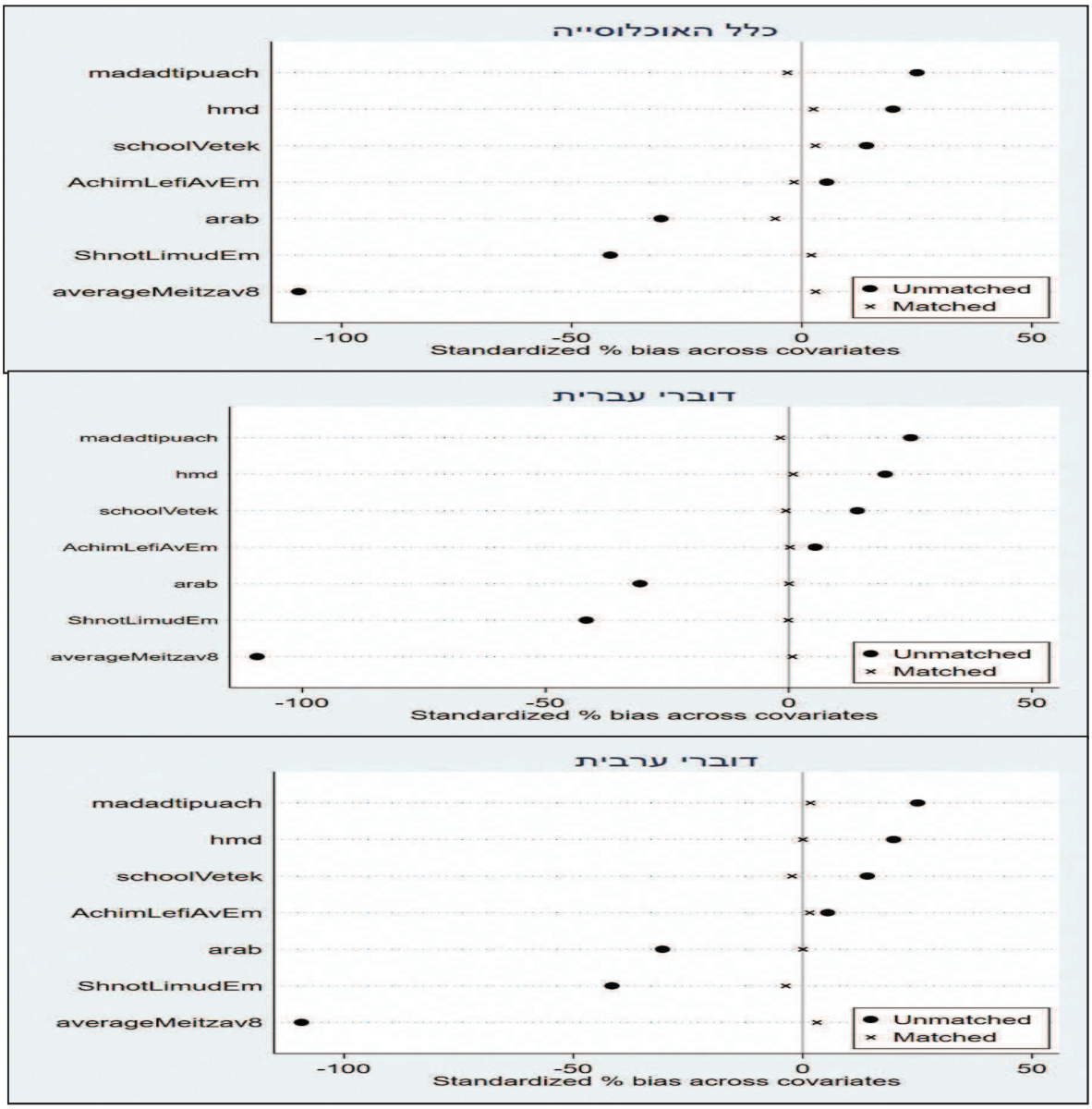
presentation of all result variables shall be divided into subgroups of Hebrew speakers and Arabic speakers.

Table 3: Statistical features of the sample, before and after adjustment

Variable	Balance test – entire sample							
	Adjustment	**Treated group**	**Peer group**	**Difference**	**% bias**	**% reduce bias**	**t**	**p>t**
		Student variables						
Meitzav test score, 8th grade	Before	32.00	56.39	-24.40	-109.3		-117.69	0
	After	32.20	31.54	0.66	3	97.3	2.78	0.005
Years of education – mother	Before	12.03	13.22	-1.19	-41.6		-66.25	0
	After	12.00	11.94	0.06	2	95.1	1.83	0.068
Number of siblings	Before	2.87	2.78	0.09	5.4		9.39	0
	After	2.75	2.78	-0.03	-1.7	68.7	-1.44	0.149
		School variables						
School cultivation grade	Before	5.68	5.10	0.58	25.1		42.25	0
	After	5.54	5.61	-0.07	-3.1	87.4	-2.75	0.006
Median teaching seniority	Before	19.75	19.19	0.56	14.1		24.51	0
	After	19.98	19.86	0.12	2.9	79.2	2.59	0.01
Percentage of students in State-Religious schools	Before	0.22	0.14	0.08	19.8		38.32	0
	After	0.19	0.18	0.01	2.5	87.4	2.02	0.043
Percentage of students in State-Arab schools	Before	0.15	0.27	-0.12	-30.6		-50.23	0
	After	0.13	0.16	-0.02	-5.8	81	-5.51	0

Figure 2 shows the different variables before and after matching. The Y axis shows the names of explanatory variables; The X axis shows the standardized bias across covariate (the difference between the groups); the full circles on the figure indicate the difference in percent between the groups before matching; The X marks indicate the difference after matching. The better the matching, the closer the values after matching to zero on the X axis. You may see that in figures for sub-groups of Hebrew speakers and Arabic speakers, most of the variables are close to the zero line, which indicates a better matching of the groups. Table 4 presents the estimation results of the probability model, after applying the matching algorithm. You may see that for the three samples, the explanatory power of the model after matching is close to zero, and the median and average bias rate between the groups are below the 5 percent threshold.

Figure 2: <u>Graphic description of the difference between groups in observed variables before and after group matching</u>



The variables are: Nurture index, State-Religious school system, school teacher seniority, number of siblings, percentage of Arabs, years of education of mother, and average 8th grade Meitzav score.

Table 4: Estimation results of the probability model, after applying the matching algorithm

	Estimation of probability model after group matching				
	Ps R2	**LR chi2**	**p>chi2**	**Mean-Bias**	**Med-Bias**
Before adjustment	0.174	14136.53	0.00	35.1	25.1
Entire sample	0.001	39.10	0.00	3	2.9
Hebrew speakers	0.000	4.88	0.559	0.6	0.6
Arabic speakers	0.001	4.39	0.495	1.8	1.6

5.2 *Impact of participation in the program on result variables*

After we have estimated the probability of participation in the program and ensured that the sample is balanced, we test below the differences between groups with regard to several result variables. The first result variable we test is the impact of attending an Mabar class on dropout rate from high school. For each result table, we present the average of the result variable for Mabar Class, and the estimator for program impact is the difference between students who attended an Mabar class and students with similar attributes who attended an ordinary class.

Results shown in Table 5 indicate that attending an Mabar class resulted in significant reduction of the dropout rate. This result is consistent: It was obtained both in the WLS estimation (after assignment of an appropriate weighting to each observation), as well in both of the matching models. The dropout rate for Mabar classes was lower by (2.9)–(3.3) percentage points, compared to students with similar attributes who attended an ordinary class. That is to say, participation in Mabar reduced the dropout rate from 8 percent, the expected value without participation in the program, to 4.7 percent for students in Mabar classes, a 40 percent reduction.[15]

Review of the impact among Hebrew speakers only shows that participation in the program results in a reduced dropout rate, but to a lesser extent than for the entire sample. According to the three models estimated, attending an Mabar class results in reduction of the dropout rate by (1.0)–(1.7) percentage points, to a dropout rate of 5.0 percentage points, compared to 6.6 expected without the program. Review of the program's impact on Arabic-speaking students shows that the dropout rate for students who participated in the program was significantly lower than for the peer group, and that the dropout rate for students who

[15] To calculate the dropout rate / matriculation eligibility rate expected without participation in the program, we must add the average of the result variable in Mabar classes to the estimator for the relevant PSM.

participated in the program was reduced by (5.2)–(6.8) percentage points compared to students with similar attributes who attended an ordinary class. The dropout rate for Arabic speakers who attended an Mabar class was 2.8 percent, compared to the expected 9 percent without participation in the program, a 70 percent reduction.

Table 5: Estimation results of variable Dropout Rate

	Dropout rate		
	WLS (1)	**Nearest Neighbors** (2)	**Kernel** (3)
Entire population Mabar average		4.7%	
Estimated program impact	-0.0294*** (0.003)	-0.0320*** (0.002)	-0.0330*** (0.003)
Hebrew speakers Mabar average		5.0%	
Estimated program impact	-0.0109*** (0.003)	-0.0155*** (0.003)	-0.0172*** (0.003)
Arabic speakers Mabar average		2.8%	
Estimated program impact	-0.0525*** (0.007)	-0.0623*** (0.005)	-0.0681*** (0.008)

Table 6 shows the estimation results of the program's impact on the matriculation eligibility rate for students who started in 10th grade. The results indicate that attending an Mabar class significantly increases the matriculation eligibility rate for all models reviewed. The matriculation eligibility rate for Mabar students was 70 percent, higher by 4.8–4.9 percentage points compared to students with similar attributes who attended an ordinary class. As with the dropout rate variable, here too the program had different impact on the subgroups: Among Hebrew speakers, the program resulted in an increase by 1.4–1.9 percentage points in the matriculation eligibility rate, but for Arabic speakers the impact was more pronounced: the WLS model resulted in an increase by 6.1 percentage points in eligibility rate, and the two matching models showed an impact of 9.1–9.9 percentage points.

Another result variable we tested was the average matriculation score for those taking the matriculation test (i.e., excluding dropouts), including bonus points, as presented in Table 7. The estimation results show that of those students who took the matriculation test, the average score for Mabar students was lower by (2.7)–(3.2) points compared to students

with similar attributes who attended an ordinary class.. The negative impact of the program on student achievement may be due to several factors: First, the difference may result from how the variable is defined, i.e., calculated excluding dropouts, which precludes reflecting the fact that in Mabar classes the dropout rate is relatively low; The program succeeds in preventing students from dropping out, but at the cost of a lower average matriculation score for the class (the percentage of students taking the test in Mabar classes is higher by one percentage point). Second, the difference may be due to a difference in the number of curriculum units studied by students in different class types and the effect of bonus points awarded for a higher number of curriculum units. In the Mabar program, students tend to have a low number of curriculum units in mandatory subjects (21 units on average, compared to 25 units in the peer group); Estimation excluding bonus points reduced the difference to 1.1 points—but the difference remained negative and significant. Third, we cannot preclude the existence of an effect of unobserved variables correlated with test achievement; This is in line with the identification strategy, where we estimated that it may result in underestimation because the peer group is stronger than we assume. Similar to previous results, here we also saw a difference in results between Hebrew speakers and Arabic speakers. Among Hebrew speakers, the average matriculation score for Mabar students was lower by 3.4–3.7 points, whereas among Arabic speakers, the WLS model resulted in a significant negative difference of 1.1 points and the two matching models resulted in nonsignificant negative estimators (0.5 points).

Table 6: Estimation results of variable Matriculation Eligibility Rate

	Matriculation eligibility rate		
	WLS	**Nearest Neighbors**	**Kernel**
	(1)	(2)	(3)
Entire population Mabar average		70.0 percent	
Estimated program impact	0.0494*** (0.007)	0.0485*** (0.005)	0.0476*** (0.006)
Hebrew speakers Mabar average		72.0 percent	
Estimated program impact	0.0194*** (0.007)	0.0138*** (0.005)	0.0134** (0.006)
Arabic speakers Mabar average		56.0 percent	
Estimated program impact	0.0609*** (0.021)	0.0914*** (0.013)	0.0985*** (0.017)

Another analysis is for subgroups within the group of Hebrew speakers—the State-Hebrew school system and the State-Religious school system—for key results see Appendix 8.3. Testing the impact of Mabar classes in State-Hebrew schools showed that participation in the program reduced the dropout rate by (1.8) percentage points and increased the matriculation eligibility rate by 1.8 percentage points compared to students with similar attributes who attended an ordinary class. Among students in State-Religious schools, we saw no significant difference that participation in the program had on reducing the dropout rate and increasing the matriculation eligibility rate.

Table 7: Estimation results of variable Average on Matriculation Exam (incl. bonus) out of those tested

	Average matriculation score for those tested		
	WLS	**Nearest Neighbors**	**Kernel**
	(1)	(2)	(3)
Entire population Mabar average		73.4	
Estimated program impact	-2.684*** (0.174)	-3.060*** (0.138)	-3.176*** (0.187)
Hebrew speakers Mabar average		73.5	
Estimated program impact	-3.442*** (0.169)	-3.615*** (0.141)	-3.672*** (0.186)
Arabic speakers Mabar average		72.4	
Estimated program impact	-1.117** (0.529)	-0.524 (0.404)	-0.521 (0.560)

Another test involved the program impact by gender. The sample was divided into boys and girls, with estimation resulted shown in an appendix table. The results indicate that among girls, the program reduced the dropout rate by (1.3)–(1.6) percentage points, whereas among the boys this rate was reduced by 4.5–5.0 percentage points. We saw a similar indication, of a stronger impact on boys, with regard to the matriculation eligibility rate as well. Girls who participated in the program increased their matriculation eligibility rate by 1.8–1.9 percentage points according to the two matching models, and by 3.9 percentage points according to the WLS model, whereas among the boys, attending Mabar classes improved the matriculation eligibility rate by 6.3–7.4 percentage points. Similar to all other results so far, for the sub-groups as well, attending an Mabar class also resulted in a decrease in average matriculation score of participants compared to students with

similar attributes who attended an ordinary class. The average matriculation score for girls who participated in the program was lower by (2.9)–(3.5) points, and for boys it was lower by (2.4)–(2.6) points compared to students with similar attributes who attended an ordinary class.

The study results so far indicate that the program may have a differential impact. This is based on the difference in impact on achievement of the weaker groups tested (Arabic speakers and boys) compared to their peers. In order to test the hypothesis whereby as the program is allocated to a weaker population, its benefit in reducing the dropout rate and increasing the matriculation eligibility rate is more pronounced, we shall estimate the model based on 3 sub-samples of students with weak background attributes and compare these estimates to the ones obtained for the general population. For the first sample we imposed a restriction whereby the number of school years of mothers of program participants and of those in the peer group classes, is no more than 12; For the other two samples, we imposed a restriction regarding student achievement in the 8th grade Meitzav test (achievement up to the median and in the bottom third of the achievement ranking). Results shown in Table 8 indicate that among students whose parents have up to 12 years of education, the program impact on reducing the dropout rate and increasing the matriculation eligibility rate was 25 percent higher compared to results in the general population (dropout rate reduced by 3.9 percentage points and eligibility rate increased by 6.2 percentage points). Review of the impact on students whose achievement in the 8th grade Meitzav test was below median (column 2) showed that the program impact was 50 percent higher compared to the general population, with the dropout rate reduced by 4.2 percentage points and the eligibility rate increased by 8.2 percentage points. Column 3 shows the program impact on students in the bottom third of achievement in the 8th grade Meitzav test. Test results show that the program had a significantly higher impact compared to all of the aforementioned samples, with the dropout rate reduced by 6.8 percentage points compared to 3.2 percentage points in the general population, and the matriculation eligibility rate increased by 14.8 percentage points compared to 4.9 percentage points in the general population.

Table 8: Testing of differential program impact

	Differential program impact		
	Parent education up to 12 years of education (1)	**Below median achievement in 8th grade Meitzav test** (2)	**Bottom third of achievement in 8th grade Meitzav test** (3)
Dropout rate			
Average for treated group	4.9 percent	5.1 percent	5.6 percent
Estimated program impact	-0.0395*** (0.003)	-0.0421*** (0.003)	-0.0685*** (0.004)
Average matriculation eligibility rate			
For treated group	67.1 percent	67.3 percent	63.4 percent
Estimated program impact	0.0624*** (0.006)	0.0820*** (0.006)	0.1480*** (0.008)
Average matriculation score for those tested			
Average for treated group	72.4	71.9	70.0
Estimated program impact	-2.546*** (0.159)	-1.792*** (0.149)	-0.294 (0.201)

5.3 *Robustness test*

The study was constructed such that all results presented so far were conservative, since the bias effect was inherently biased toward achievement in the control group. To reduce this bias and to test the robustness of the results, we introduce changes to the sample, so as to ensure that the assumptions underlying the study are not overly powerful. The identification strategy in this study relied on the following:

1) Excess demand for the program in each year;

2) Excess students directed to ordinary classes, based on a decision made by the professional team at the school.

The first test includes estimation of the model for schools where an Mabar **class was first opened during the study period (2010–2015).**[16] Estimation of this sample ensures that prior to the class being opened there was excess demand for the program, as the school had students who met the preconditions but had no Mabar class available. This change to the sample reduces the number of observations from 140,000 students (of which 13,500 in Mabar classes) to 16,000 students (of which 1,300 in Mabar classes). The disadvantage of this estimate is that the school sample no longer represents the general population, with over-representation of Arabic-speaking schools and under-representation of State-

[16] In fact, the class was opened in 2011–2015, and in 2010 there were no classes in any of the schools.

Religious schools. Another difference is that schools included in this sample are significantly weaker than in the general one.

The estimation results (Table 9) for the dropout rate variable show that participation in the program reduced the dropout rate by 4.7–7.5 percentage points—a significantly higher impact than for the estimate based on the entire sample—(2.9)–(3.3) (Table 5). When testing the impact on matriculation eligibility rate, we saw different results in the different models. In the two matching models, we found that participation in the program increased the matriculation eligibility rate by 16.5 percentage points, from 52 percent eligibility in the control group to 68 percent in the treated group. However, the OLS model, without accounting for the selection issue, did not indicate a significant impact due to participation in the program, possibly due to the inherent bias in this estimation. As for the impact on average matriculation score, we also saw mixed results: The OLS model showed that participation in the program resulted in a decrease by (3.2) points in the average matriculation score, while the matching models showed a non-significant positive estimator.

The objective of the second test is to reduce the bias caused by intervention of the professional team in assignment of students with similar attributes to Mabar classes or to ordinary classes. Based on interviews with teams of Shachar classes at the Ministry of Education, it would appear that when two students meet the conditions for admission but there is excess demand, assignment to classes is based on recommendation of the professional team, with the better students assigned to ordinary classes (for more information see Section 4.1 and Appendix 8.1), it is reasonable to assume that assignment was also based on unobserved variables of each individual. In order to reduce the impact of the bias due to unobserved variables of each individual ("True econometric selection bias"), we shall estimate **the program impact for a sample consisting exclusively of three-year schools (high school only)**. At these schools, the professional team at the school is less familiar with the students and their abilities than in six-year schools (combined middle school and high school), since the team is different when moving to the high school, compared to continuous schooling at the six-year schools. Consequently, we expect that assignment of excess students to ordinary classes would have a weaker correlation with unobserved variables of each individual, and therefore the estimated program impact should be stronger than in the model estimated based on the entire sample (reducing the B3 bias component).

Table 9: Estimated program impact in schools where a Mabar class was first opened during the study period

	Schools in which Mabar class was opened after 2010		
	OLS (1)	**Nearest Neighbors** (2)	**Kernel** (3)
Dropout rate Mabar average		4.2 percent	
Estimated program impact	-0.0467*** (0.011)	-0.0720*** (0.008)	-0.0751*** (0,011)
Matriculation eligibility rate Mabar average		68.5 percent	
Estimated program impact	0.0320 (0.024)	0.1646*** (0.016)	0.1734*** (0.020)
Average matriculation score for those tested Mabar average		74.0	
Estimated program impact	-3.187*** (0.616)	0.0887 (0.472)	0.1531 (0.650)

The estimation results are presented in Table 10. Program impact on the dropout rate shows that participation in the program reduces the dropout rate by (4.2)–(4.6) percentage points—higher by 1.2 percentage points than estimation in the overall sample. Review of the impact of participation in the program on the matriculation eligibility rate shows an increase by 5.5–8.4 percentage points in the matriculation eligibility rate. Differences between estimators in the sample of three-year schools and in the overall sample show that the selection by the professional team impacts the estimation results. When testing the impact of participation in the program on the average matriculation score for those taking the test, we see that participation in the program decreased this average by (2.7)–(2.9) points compared to students with similar attributes who attended an ordinary class.

Table 10: Estimation of program impact at three-year schools only

	Three-year schools only		
	OLS (1)	**Nearest Neighbors** (2)	**Kernel** (3)
Dropout rate Mabar average		4.3 percent	
Estimated program impact	-0.0422*** (0.005)	-0.0459*** (0.004)	-0.0456*** (0.006)
Average matriculation eligibility rate In Mabar		65.7 percent	
Estimated program impact	0.0554*** (0.015)	0.0821*** (0.009)	0.0840*** (0.012)
Average matriculation score for those tested In Mabar		73.3	
Estimated program impact	-2.917*** (0.403)	-2.751*** (0.259)	-2.724*** (0.367)

We also tested the impact of opening an Mabar class (for the first time) on the average school achievement. This test used the OLS model with a fixed effect for school attributes and all other variables being the average of individual attributes when students were in 10^{th} grade. In this estimation we defined two types of treatment variables: 1) A binary variable assigned the value 1 if an Mabar class existed in the age group, or 0 otherwise. 2) Percentage of Mabar students out of total students in Mabar and ordinary classes at the school (14 percent on average for the entire period, up to 20 percent for schools where a class was first opened). We estimated this based on two different specifications of the result variable (linear and logarithmic) due to the differential program impact, as observed in previous sections. Appendix 8.4 shows a table of descriptive statistics for this sample, which includes weaker schools than in the overall sample, which is reflected in a lower Meitzav score, fewer years of education of mothers, a higher cultivation grade and a higher percentage of Arabic-speaking schools.

Table 11 shows the key results for the school dropout rate variable, by specification of the result and treatment variables. Columns 1 and 3 show the impact of **opening an** Mabar **class** on the school dropout rate. The linear model shows no significant impact (but nearly significant at the 10 percent significance level), while the semi-logarithmic model shows that opening an Mabar class results in a (24 percent) reduction in the school dropout rate—

equivalent to a (1.1) percentage points decrease in dropout.[17] Assuming that this reduction in the dropout rate was exclusively due to the impact on Mabar students (with no spillover), then participation in the program reduces the dropout rate by 7.8 percentage points (dividing the dropout rate in percentage points by the percentage of Mabar students in the sample, which is 14 percent)—similar to the matching results. Another test involved the impact of percentage of Mabar students on the school dropout rate (columns 2 and 4), and the meaning of this estimator is that a 1 percent increase in the percentage of Mabar students resulted in a 1 percent decrease in the school dropout rate. Here, too, we saw no significant impact in the linear model, with a significant impact indicated by the semi-logarithmic model: A 1 percent increase in the percentage of Mabar students resulted in a (0.7 percent) decrease in the school dropout rate. Results shown in Table 11 indicate that in the linear model, we could not identify any program impact on the dropout rate (although for both estimates, the significance level is nearly 10 percent). A possible explanation of this result lies in the differential program impact (as observed in previous sections), which results in a stronger impact when comparing weaker groups (by school level, gender, and so on).

Table 11: Estimated program impact on school dropout rate

	Impact on school dropout rate			
	Linear model		**Semi-logarithmic model**	
	(1)	(2)	(3)	(4)
Mabar class first opened	-0.029 (0.020)		-0.271** (0.129)	
Percentage of Mabar students		-0.228 (0.169)		-1.285* (0.738)
Fixed effect	V	V	V	V
Number of observations	266	265	266	265
R-squared	0.088	0.246	0.152	0.187
Number of schools	45	45	45	45

Table 12 shows the program impact on the school matriculation eligibility rate. Columns 1 and 3 show the impact of opening an Mabar class on the school matriculation eligibility rate—an increase by 4–6 percent compared to school achievement in years prior to opening the class, i.e., an improvement by 3.1–4.7 percentage points. In columns 2 and 4 we see the impact of higher percentage of Mabar students on the school matriculation eligibility

[17] Since the result variable is semi-logarithmic, the meaning of the estimator is: $g = \exp(\beta) - 1$

rate: A 10 percent increase in the percentage of Mabar students results in a 1.5–2.5 percent increase in the school matriculation eligibility rate. Assuming the impact is entirely on Mabar students, this estimator shows a 17 percent improvement.

Table 12: Estimated program impact on matriculation eligibility rate

	Impact on school matriculation eligibility rate			
	Linear model		Semi-logarithmic model	
	(1)	(2)	(3)	(4)
Mabar class first opened	0.042*** (0.009)		0.063*** (0.015)	
Percentage of Mabar students		0.146*** (0.047)		0.230*** (0.082)
Fixed effect	V	V	V	V
Number of observations	266	265	266	265
R-squared	0.188	0.207	0.176	0.197
Number of schools	45	45	45	45

Another test concerned the school average matriculation score (Table 13). Estimation at the individual level indicates that the average matriculation score among program participants was lower than for the peer group. Estimation at the school level indicates that opening the first Mabar class results in a significant increase by 0.5 points in the school average matriculation score. This result was obtained using both the linear model and the semi-logarithmic model. Testing the impact of higher percentage of Mabar students resulted in a nonsignificant impact in the two models used. The difference in results between estimation at the individual level and estimation at the school level supports the possibility that opening an Mabar class has a spillover impact on students in the ordinary classes, whether due to the weakest students being removed and assigned to a designated class, or due to training provided to the entire teaching staff.

Table 13: Estimated program impact on average school grade on matriculation diploma

	Impact on school matriculation average			
	Linear model		Semi-logarithmic model	
	(1)	(2)	(3)	(4)
Mabar class first opened	0.543** (0.260)		0.006* (0.003)	
Percentage of Mabar students		0.100 (1.049)		-0.001 (0.013)
Constant effect	V	V	V	V
Number of observations	264	263	264	263
R-squared	0.176	0.167	0.175	0.169
Number of schools	45	45	45	45

All of the robustness tests described indicate a significant underestimation of results estimated across the entire sample. In the sample designed to address the assumption of lack of appropriate peer group in the ordinary classes, a significantly higher impact was shown than the overall results—the dropout rate declined by 7.2 percentage points, compared to 3.3 percentage points, and matriculation eligibility rate improved by 16 percent, compared to 5 percent in the general model. This impact is partially due to the fact that the program was provided to weaker schools than those included in the general sample; based on an interview with Shachar staff, these are mostly "red" schools, which are not representative of the general population.[18] The second test, designed to reduce the impact of unobserved attributes in the control group (due to impact of the school staff), showed that the program reduced the dropout rate by 5 percent and increased the matriculation eligibility rate by 8 percent—this is a more reasonable estimate and we shall use it to evaluate the program benefit.

6. Estimated program cost/benefit

Study results show that every year, the Mabar program prevents 400–550 students from dropping out of school and helps another 600–1,000 students to obtain a matriculation diploma.[19] This is at an incremental cost of NIS 350 million per year. Thus, the marginal cost of improving the achievement of 1,000–1,550 students is roughly estimated at NIS

[18] Schools that require improvement, based on definitions of the Ministry of Education; For more information see: "Red Schools—Policy and Research Overview" Dr. Odette Sela, April 29, 2015.

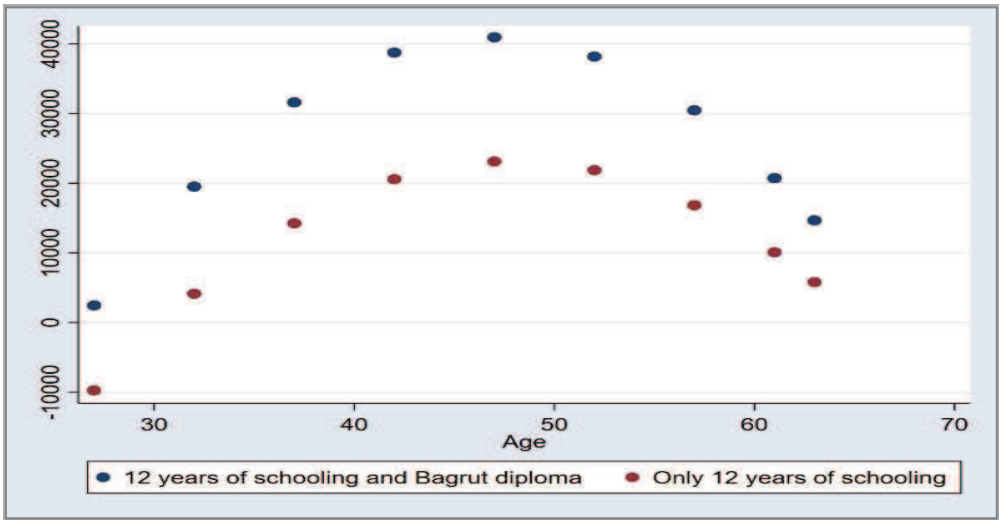
[19] The difference between the estimates is due to difference in results between Tables 5 and 6 and Table 10.

225,000–350,000 per student. However, in order to evaluate whether this investment is worthwhile, we must understand the benefits of preventing school dropouts and increasing the matriculation eligibility rate. No comprehensive study has been conducted in Israel with regard to the economic cost and impact of dropping out of school and of finishing high school without a matriculation diploma, on the welfare of individuals in their adult life. In the US, studies found that the earnings difference between employees who dropped out of high school and employees who finished high school (at least) over their life time was estimated at $630,000, and the economic cost of every school dropout was estimated at $240,000—the result of higher welfare, healthcare and incarceration expenses (Chapman, et al. 2011).

In absence of any research that quantifies the benefit of finishing 12 years of education, and of graduating from high school with a matriculation diploma, we shall try to estimate, using the Household Expenditure survey of the Central Bureau of Statistics, the difference in earnings between individuals with such education levels. In order to estimate the impact of variance in education levels on future earnings, we estimated a Mincer earnings function for a sample of non-ultra-Orthodox Jews and Arabs with 9-12 years of education. We distinguished between dropouts (fewer than 12 years of education), those with 12 years of education but no matriculation diploma, and those with 12 years of education and a matriculation diploma.

The objective of this test is to estimate the difference in earnings between these education groups in the primary working ages and, in particular, to test the earnings difference between students who completed 12 years of education and students who dropped out of high school, and between students who completed 12 years of education with a matriculation diploma and students who students who completed 12 years of education without one. The key results are shown in Figure 3 and in Appendix 8.4. The X axis shows the employee age and the Y axis shows the annual difference in earnings compared to employees who did not complete 12 years of education.

Figure 3: Estimated annual wage difference, in NIS, between employee with 12 school years completed and employee with 12 school years completed and matriculation diploma vs. employee who dropped out of high school



Using the regression coefficients, we estimated the difference in earnings for an employee in the employment market over the working years (ages 25–64). The lifetime difference in earnings between an employee who dropped out after 9–11 years of education and an employee who completed 12 years of education is estimated at NIS 491,000; the difference in earnings between an employee who obtained a matriculation diploma and an employee who merely completed 12 years of education is estimated at NIS 607,000 (a difference of NIS 1.1 million compared to an employee who dropped out). Furthermore, many studies around the world indicate that education is negatively correlated with welfare payments and healthcare system costs. Another factor not included here is the cost of crime. According to data from the Israeli Prison Service for 2018 and data from the Central Bureau of Statistics, 0.6 percent of those who did not complete 12 years of education are incarcerated, compared to 0.04 percent of those who completed 12 years of education and 0.02 percent of those who have at least a matriculation diploma.[20] Given the difficulty of quantifying these costs in the absence of comprehensive research in Israel, we shall estimate that the same ratio of earnings to welfare payments found in the US (25 percent) is also true in Israel, hence the benefit from preventing dropout is estimated at NIS 614,000 and the benefit from obtaining a matriculation diploma is estimated at NIS 759,000 (NIS 1.375 million compared to a school dropout).

[20] Israeli Prison Service report for 2018, National Prison Service. Survey of household expenses and income by the Central Bureau of Statistics, 2018.

However, since these are future amounts calculated over 40 years, we should discount them to their present value. The appropriate discount rate is in dispute, and is affected by various assumptions. In Table 13 we present a sensitivity analysis for various discount rates from 1 percent to 3 percent. The objective of these low discount rates is to represent other impacts over time, due to improved education, such as the inter-generation benefit which, under certain scenarios, is expected to result in infinite benefit of the program[21] (Behrman & Rosenzweig, 2002; Currie & Moretti, 2003; Black & Devereux, 2005). The high discount rates reflect the risk associated with long-term investment in a weak population, which may be significantly more susceptible to exogenous shocks in the employment market than other populations. The 3 percent discount rate is higher than the State's cost of borrowing, estimated at 2.2 percent on average in the study period, and higher than the current cost of long-term debt (2 percent). The study results show that based on the estimated reduction in dropouts and increase in matriculation eligibility, based on an estimated 1,000 students—the estimated benefit of the program is NIS 307–527 million per year, and based on an estimated 1,500 students is it NIS 485–825 million per year. Sub-analysis for the Arabic-speaking population shows that while the overall program cost for this population is estimated at NIS 52 million, the program benefit is estimated at NIS 92–142 million, depending on the discount rate applied (Appendix 8.5).

This analysis merely provides a rough estimate of the program benefit, since it relies on the assumption that students who obtain a matriculation diploma thanks to the program do not go on to higher education, which artificially reduces the program benefit.[22] In contrast, Mabar students may have better skills than those of the average dropout student and inferior to those of the average matriculation diploma holder (comparing the strongest dropouts to the weakest matriculation diploma holders), in which case, the regression estimates calculated for the entire population may result in over-estimation of the program benefit. We would recommend subsequent studies to test the long-term impact of reduced

[21] If the program resulted in improved education of an individual in generation t, the education of their descendants in future generations t+1 through t+n would be impacted by this improvement in education level, assuming that return on education is higher than zero.

[22] Based on administrative data available to us, 30 percent of students in MBR classes went on to higher education by ages 27-28.

[22] Given that the program resulted in improved education of an individual in generation t, the education of their descendants in future generations t+1 through t+n would be impacted by this improvement in education level due to effect of the program, assuming that return on education is higher than zero.

dropout rate and increased matriculation eligibility rate on increase in work inputs, education and reduced cost of welfare and crime.

Table 13: Sensitivity analysis of estimated program benefit using different discount rates

	Sensitivity analysis of estimated program benefit		
Discount rate	3 percent	2 percent	1 percent
Prevent dropout	185,104	221,131	315,023
Matriculation diploma	286,280	389,867	495,428
Estimate per 1,000 (NIS million)	307	400	527
Estimate per 1,500 (NIS million)	**485**	**629**	**826**

* Estimate per 1,000 was calculated based on improved achievement of 1,000 students per year, whereas the estimate per 1,500 was calculated based on improved achievement of 1,500 students per year.

7. Summary

In this study we tested the effectiveness of Mabar classes, a program designed to reduce the school dropout rate and to increase the matriculation eligibility rate among students from a low socioeconomic background. This program has been in operation in high schools since the mid 1990s, at an incremental annual cost of NIS 350 million. Despite the long period in which this program has been in operation and its important objectives, its effectiveness has not been tested to date. That is primarily due to methodology challenges, and in particular the selection bias of students for this program, which is expected to result in bias of its estimated benefit.

By analysis of data for all students who started high school (10^{th} grade) in non-vocational schools in 2010–2015, we tested the effectiveness of the Mabar program in reducing the dropout rate and increasing the matriculation eligibility rate, as well as its impact on the average matriculation diploma scores of those tested (including bonus points). In order to test the program effectiveness, we took advantage of the fact that in each year there was excess demand for the program, and the gradual opening of additional classes in high schools. Using the unique criteria for admission to the program, we estimated the probability of students to be included in the program, and based on this probability we created a peer group of students in ordinary classes with attributes similar to those of students in treated classes.

The descriptive statistics indicated that the program is allocated to students with significantly weaker attributes than those of students in ordinary classes: both the socioeconomic background and the scholastic achievement of these students are significantly lower, and by a large margin, than those of their peers in ordinary classes. The descriptive statistics indicates that the program was allocated unequally to the different sectors and school systems—under-allocation to the Arab sector and over-allocation to the State-Religious school system.

The study found that, subject to strict assumptions, the program resulted in a decrease by 3.2 percentage points in the dropout rate among students who participated in the program, compared to the peer group in ordinary classes—or 40 percent. We found a stronger impact in Arabic-speaking schools. In these schools, the dropout rate decreased by 6.2 percentage points, and among boys—by 5.0 percentage points. A lesser impact was found among Hebrew speakers, and in the State-Religious school system the program was found to have no impact.

The study results show that participation in the program increased the matriculation eligibility rate by 4.9 percentage points (a 7.5 percent improvement). A stronger impact was found among Arabic speakers—improvement by 9 percentage points (a 20 percent improvement), compared to 1.4 percentage points among Hebrew speakers and no impact in the State-Religious school system. The average matriculation score of those taking part in the program is relatively low; this may be because the program is focused on bringing as many students as possible to cross the threshold required to obtain a matriculation diploma (for example, by precluding students from taking tests on a large number of curriculum units), which results in a change in composition of students compared to the peer group. In contrast, we found that opening an Mabar class results in improvement in the school average matriculation score—possibly due to the weaker students being removed from ordinary classes.

The study shows that the program has a differential impact: As the program is allocated to a weaker population, its benefit in reducing the dropout rate and increasing the matriculation eligibility rate is more pronounced. Students who are admitted with the lowest Meitzav scores succeed far more than students with the same Meitzav scores who study in ordinary classes. We also observed a strong program impact on Arabic-speaking students and on students whose parents have up to 12 years of education.

The results of this work are conservative, since they were obtained subject to strict assumptions that result in under-estimation of program impact. In order to test robustness, we tested the program impact using two sub-samples, which allowed us to estimate this impact more cleanly. In the first part, we estimated the program impact on schools where a first Mabar class was opened. We found that the dropout rate was reduced by 7 percentage points and the matriculation eligibility rate was increased by 16 percentage points. Another test looked at three-year high schools. In those schools, the program impact was estimated at the dropout rate was reduced by 8 percentage points and the matriculation eligibility rate was increased by 8 percentage points.

Another result is the lack of program impact in the State-Religious school system, which receives the highest allocation within this program (in terms of participants as percentage of total students in the school system). We believe that this is due to relatively strong students being assigned to Mabar classes, and being compared to relatively strong students in ordinary classes (since the peer group is made up of students with similar attributes to those of students in the treated group). It may also be the case that in this school system, with extensive allocation of Mabar classes, there was no excess demand and therefore no weak students were assigned to ordinary classes.

References

Abadie, A. & G. W. Imbens (2002). *Simple and bias-corrected matching estimators for average treatment effects* (No. t0283). National Bureau of Economic Research.

Allensworth, E. M. & J. Q. Easton (2007). What Matters for Staying On-Track and Graduating in Chicago Public High Schools: A Close Look at Course Grades, Failures, and Attendance in the Freshman Year. Research Report. Consortium on Chicago School Research.

Angrist, J. D. & V. Lavy (1999). Using Maimonides' rule to estimate the effect of class size on scholastic achievement. The Quarterly Journal of Economics, 114(2), 533–575.

Altinok, N. & G. Kingdon (2012). New evidence on class size effects: A pupil fixed effects approach. Oxford Bulletin of Economics and Statistics, 74(2), 203–234.

Balfanz, R., L. Herzog & D. J. Mac Iver (2007). Preventing student disengagement and keeping students on the graduation path in urban middle-grades schools: Early identification and effective interventions. Educational Psychologist, 42(4), 223–235.

Banerjee, A., R. Banerji, J. Berry, E. Duflo, H. Kannan, S. Mukherji, M. Shotland, & M. Walton (2016). Mainstreaming an effective intervention: Evidence from randomized evaluations of "Teaching at the Right Level" in India (No. w22746). National Bureau of Economic Research.

Banerjee, A. V., R. Banerji , E. Duflo, R. Glennerster & S. Khemani (2010). "Pitfalls of participatory programs: Evidence from a randomized evaluation in education in India*", American Economic Journal: Economic Policy*, 2(1), 1–30.

Banerjee, A. V., S. Cole, E. Duflo & L. Linden (2007). "Remedying education: Evidence from two randomized experiments in India", The Quarterly Journal of Economics, 122(3), 1235–1264.

Behrman, J. R. & M. R. Rosenzweig (2002). "Does increasing women's schooling raise the schooling of the next generation?", American Economic Review, 92(1), 323–334.

Belfield, C. R. & H. M. Levin. (Eds.). (2007). *The price we pay: Economic and social consequences of inadequate education*, Brookings Institution Press.

Bessho, S. I., H. Noguchi, A. Kawamura, R. Tanaka & K. Ushijima, (2019). "Evaluating remedial education in elementary schools: Administrative data from a municipality in Japan", *Japan and the World Economy*, 50, 36–46.

Black, S. E., P. J. Devereux & K. G. Salvanes (2005). "Why the apple doesn't fall far: Understanding intergenerational transmission of human capital". American Economic Review, 95(1), 437–449.

Bowers, A. J., R. Sprott & S. A. Taff (2012). “Do we know who will drop out? A review of the predictors of dropping out of high school: Precision, sensitivity, and specificity”, *The High School Journal*, 77–100.

Caprara, G. V., M. Vecchione, G. Alessandri, M. Gerbino & C. Barbaranelli (2011). “The contribution of personality traits and self-efficacy beliefs to academic achievement: A longitudinal study”, *British Journal of Educational Psychology*, 81(1), 78–96.

Currie, J. & E. Moretti (2003). “Mother's education and the intergenerational transmission of human capital: Evidence from college openings”, The Quarterly Journal of Economics, 118(4), 1495–1532

Dalton, B., E. Glennie & S. J. Ingels (2009). Late High School Dropouts: Characteristics, Experiences, and Changes Across Cohorts. Descriptive Analysis Report. NCES 2009-307. National Center for Education Statistics.

Dovrin N. (2015). Equal Opportunity in Education – Demographic and Socioeconomic Barriers. Series of working papers, Chief Scientist Division, Central Bureau of Statistics. [in Hebrew]

Eren, O., B. Depew & S. Barnes (2017). “Test-based promotion policies, dropping out, and juvenile crime”, *Journal of Public Economics*, 153, 9–31.

Freeman, J. & B. Simonsen (2015). “Examining the impact of policy and practice interventions on high school dropout and school completion rates: A systematic review of the literature”, *Review of Educational Research*, 85(2), 205–248.

Heckman, J. J., H. Ichimura & P. Todd (1998). “Matching as an econometric evaluation estimator”, The Review of Economic Studies, 65(2), 261–294.

Heckman, J. J., J. Smith & N. Clements (1997). “Making the most out of programme evaluations and social experiments: Accounting for heterogeneity in programme impacts”, *The Review of Economic Studies*, 64(4), 487–535.

Hefner-Packer, R. (1990*). Alternative education programs: A prescription for success*. College of Education, University of Georgia.

Hosmer Jr, D. W., Lemeshow, S., & Sturdivant, R. X. (2013) Applied logistic regression (Vol. 398). John Wiley & Sons.

Jensen, E. (2013). “How poverty affects classroom engagement”, Educational Leadership, 70(8), 24–30.

Kohler, U. & F. Kreuter (2005). *Data analysis using Stata*. Stata press.

Lavy, V., A. Kott & G. Rachkovski (2018). Does Remedial Education at Late Childhood Pay Off After All? Long-Run Consequences for University Schooling, Labor Market Outcomes and Inter-Generational Mobility (No. w25332). National Bureau of Economic Research.

Lavy, V. & A. Schlosser (2005). "Targeted Remedial Education for Underperforming Teenagers: Costs and Benefits", *Journal of Labor Economics*, 23(4), 839–874.

Lee, N. & J. Wright, (2011). Off the map? The geography of NEETs. Work Foundation. Unpublished report.

Lleras-Muney, A. (2005). "The relationship between education and adult mortality in the United States". *The Review of Economic Studies*, 72(1), 189–221.

Leuven, E. & B. Sianesi (2003). PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing.

Mosteller, F. (1995). The Tennessee study of class size in the early school grades. The future of children, 113–127.

OECD (2019). PISA 2018 Results (Volume I): What Students Know and Can Do, PISA, OECD Publishing, Paris.

Quinn, M. M., J. M. Poirier, S. E., Faller, Gable, & S. W. Tonelson (2006). "An examination of school climate in effective alternative programs. Preventing School Failure, *Alternative Education for Children and Youth*, 51(1), 11–17.

REMA, (2019). Mabar program (Regular Matriculation Track) – Evaluation Study, 2017. [in Hebrew]

Rivkin, S. G., E. A. Hanushek & J. F. Kain (2005). "Teachers, schools, and academic achievement", *Econometrica*, 73(2), 417–458.

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

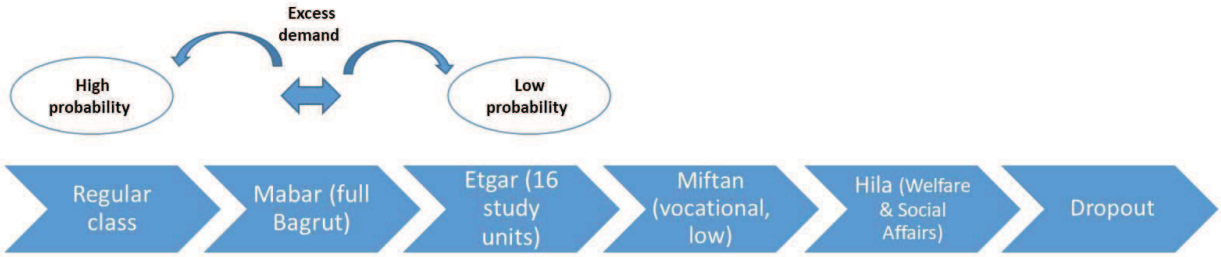
Silver, D., M. Saunders & E. Zarate (2008). What factors predict high school graduation in the Los Angeles Unified School District. Policy Brief, 14.

Zhang, Z., H. J. Kim, G. Lonjon & Y. Zhu (2019). "Balance diagnostics after propensity score matching", Annals of Translational Medicine, 7(1)

8. Appendices

8.1 *Description of assignment method to different class types for 10th grade student*

- A 10th grade student may be admitted to one of multiple class types, depending on their achievement in 9th grade and as recommended by the professional staff. The classes differ by target population and teaching objectives.
- The lion's share of students are assigned to ordinary classes, but as for some students—the professional team at the school considers that assigning them to an ordinary class may result in their not fully realizing their potential, and such students are assigned to alternative classes.
- Alternative classes differ by teaching objectives, and the objective of an Mabar class is to have the students obtain a full matriculation diploma, with other classes having lesser teaching objectives, given the students' capabilities. Thus, for example, the objective of an Mabar class is to obtain a full matriculation diploma, while the objective of an ETGAR class is to achieve 16 curriculum units.
- In case of "excess demand" for any class type, borderline students are diverted to a different class type. Mabar students are closest in attributes to students in ordinary classes, and such diversion allows us to create a peer group consisting of students in ordinary classes.



2

8.2 *Results of balance tests by sub-groups of Hebrew speakers and Arabic speakers*

8.2.1 Balance table – Hebrew speakers

Variable	Balance table – Hebrew speakers							
	Suitability	**Treated group**	**Peer group**	**Difference**	**Percent bias**	**bias**	**t**	**p>t**
Meitzav test score in 8th grade	**Student variables**							
	Before	32.00	56.39	-24.40	-109.3		-117.69	0
	After	31.07	30.92	0.15	0.7	99.4	0.63	0.532
Years of education – mother	Before	12.03	13.22	-1.19	-41.6		-66.25	0
	After	12.28	12.28	-0.00	-0.1	99.7	-0.12	0.903
Number of siblings	Before	5.68	5.10	0.58	25.1		42.25	0
	After	5.17	5.21	-0.04	-1.8	92.9	-1.54	0.122
School cultivation grade	**School variables**							
	Before	19.75	19.19	0.56	14.1		24.51	0
	After	20.51	20.53	-0.03	-0.6	95.5	-0.6	0.552
Median teaching seniority	Before	2.87	2.78	0.09	5.4		9.39	0
	After	2.57	2.57	0.00	0.2	96.3	0.17	0.863
Percentage in State-Religious schools	Before	0.22	0.14	0.08	19.8		38.32	0
	After	0.22	0.21	0.00	0.9	95.6	0.62	0.534
Percentage in State-Arab schools	Before	0.15	0.27	-0.12	-30.6		-50.23	0
	After	–	–	–	0	100	.	.

8.2.2 Balance table – Arabic speakers

Variable	Balance table – Arabic speakers							
	Suitability	**Treated group**	**Peer group**	**Difference**	**Percent bias**	**bias**	**t**	**p>t**
Meitzav test score in 8th grade	**Student variables**							
	Before	32.00	56.39	-24.40	-109.3		-117.69	0
	After	39.56	38.87	0.68	3.1	97.2	1	0.318
Years of education – mother	Before	12.03	13.22	-1.19	-41.6		-66.25	0
	After	10.14	10.25	-0.11	-3.8	91	-1.35	0.177
Number of siblings	Before	5.68	5.10	0.58	25.1		42.25	0
	After	7.95	7.92	0.04	1.6	93.4	1.18	0.237
School cultivation grade	**School variables**							
	Before	19.75	19.19	0.56	14.1		24.51	0
	After	16.55	16.64	-0.09	-2.4	83.2	-0.7	0.485
Median teaching seniority	Before	2.87	2.78	0.09	5.4		9.39	0
	After	3.85	3.82	0.02	1.5	72.9	0.42	0.677
Percentage in State-Religious schools	Before	0.22	0.14	0.08	19.8		38.32	0
	After	–	–	–	0	100	.	.
Percentage in State-Arab schools	Before	0.15	0.27	-0.12	-30.6		-50.23	0
	After	1.00	1.00	–	0	100	.	.

8.3 Further study results

8.3.1 State-Hebrew school system

	In State-Hebrew school system		
	OLS (1)	**Nearest Neighbors** (2)	**Kernel** (3)
Dropout rate Average for treated group		5.1 percent	
Estimated program impact	-0.0138*** (0.004)	-0.0178*** (0.003)	-0.0194*** (0.004)
Average matriculation eligibility rate For treated group		71.0 percent	
Estimated program impact	0.0256*** (0.008)	0.0185*** (0.006)	0.0193*** (0.007)
Average matriculation score for those tested For treated group		72.2	
Estimated program impact	-3.952*** (0.191)	-4.1424*** (0.150)	-4.0966*** (0.203)

8.3.2 State-Religious school system

	In State-Religious school system		
	OLS (1)	**Nearest Neighbors** (2)	**Kernel** (3)
Dropout rate Average for treated group		4.7 percent	
Estimated program impact	0.0068 (0.006)	-0.0056 (0.005)	-0.0062 (0.007)
Average matriculation eligibility rate For treated group		73.9 percent	
Estimated program impact	-0.0243** (0.012)	-0.0154 (0.011)	-0.0165 (0.014)
Average matriculation score for those tested For treated group		78.3	
Estimated program impact	-2.945*** (0.354)	-1.610*** (0.343)	-1.977*** (0.434)

8.3.3 Program impact by gender

	Girls			Boys		
	OLS (1)	**Nearest Neighbors** (2)	**Kernel** (3)	**OLS** (4)	**Nearest Neighbors** (5)	**Kernel** (6)
Dropout rate Mabar average		3.8 percent			6.0 percent	
Estimated program impact	-0.0135*** (0.003)	-0.0153*** (0.003)	-0.0158*** (0.003)	-0.0447*** (0.005)	-0.0497*** (0.004)	-0.0502*** (0.005)
Average matriculation eligibility rate From Mabar		73.5 percent			64.5 percent	
Estimated program impact	0.0385*** (0.008)	0.0178*** (0.006)	0.0194*** (0.008)	0.0633*** (0.011)	0.0724*** (0.008)	0.0740*** (0.001)
Average matriculation score for those tested From Mabar		74.5			71.9	
Estimated program impact	-2.942*** (0.228)	-3.548*** (0.186)	-3.550*** (0.248)	-2.394*** (0.261)	-2.645*** (0.203)	-2.604*** (0.275)

8.4 Descriptive statistics for sample of schools where a Mabar class was first opened in the study period

Average of school averages where a Mabar class was first opened after 2010		
	Average	**Standard deviation**
Meitzav test, 8th grade	51.23	8.62
Years of education – mother	12.34	2.06
Siblings	3.15	1.04
Boys	0.52	0.19
School cultivation grade	5.65	2.53
Teacher seniority (school average)	17.94	3.64
Arab school	0.44	0.50
State-Religious school	0.16	0.36
Outlying areas	0.33	0.47
Dropout rate	0.05	0.10
Eligibility rate	0.74	0.18
Average matriculation diploma score	81.86	6.77
Years of education	11.88	0.20
Percentage of Mabar students – overall	0.14	0.20
Percentage of Mabar students – after class was opened	0.20	0.22

8.5 Results of Mincer regression to estimate difference in earnings

Estimation model:

$$W_i = \alpha_0 + D_{Gender} + D_{\text{Aravi}} + D_{12} + D_{12+\text{B}} + D_{12} * Age + D_{12} * Age^2 + D_{12+\text{B}} * Age + D_{12+\text{B}} * Age^2 + year$$

Where:

W_i – *Total earnings of individual i from work as salaried employee and being self-employed.*

D_{Gender} – *Dummy variable: 1 if individual is male, 0 otherwise.*

D_{Aravi} – *Dummy variable: 1 if individual is Arab, 0 otherwise.*

D_{12} – *Dummy variable: 1 if individual has completed 12 years of education and has no matriculation diploma, 0 otherwise.*

$D_{12+\text{B}}$ – *Dummy variable: 1 if individual has completed 12 years of education and has obtained a matriculation diploma, 0 otherwise.*

Age – *Variable of age and age squared.*

$year$ – *Dummy variable for year of survey.*

Estimation results:

	Total earnings from work as salaried employee and being self-employed	
	Primary working ages	**Entire sample**
	1	2
	4,971.801***	2,959.978***
	(123.106)	(73.182)
D_{Aravi}	-3,030.602***	-1,340.507***
	(141.379)	(85.701)
$D_{12+\text{B}}$	-14,601.485***	-12,711.536***
	(1,573.643)	(365.774)
D_{12}	-12,516.219***	-10,405.073***
	(1,669.572)	(365.284)
$D_{12+\text{B}} * Age$	771.287***	799.219***
	(75.586)	(18.748)
$D_{12+\text{B}} * Age^2$	-8.256***	-8.395***
	(0.862)	(0.209)
$D_{12} * Age$	603.795***	647.014***
	(78.049)	(17.579)
$D_{12} * Age^2$	-6.308***	-6.802***
	(0.873)	(0.190)
_Iyear_1	315.646**	202.396**
	(148.529)	(88.557)
	829.792***	515.640***
	(150.455)	(89.781)
Constant	3,193.743***	838.720***
	(171.682)	(93.987)
Observations	12,811	25,964
R-squared	0.165	0.194

	Program benefit for Arabic speakers		
	3 percent	2 percent	1 percent
Prevent dropout	131,125	168,728	219,430
Matriculation diploma	116,166	140,316	169,197
Benefit from increased earnings (NIS in millions)	73	91	114
Program benefit (NIS in millions)	92	114	142