ONLINE APPENDIX

Short- and Long-Term Effects of Universal Preschool: Evidence from the Arab Population in Israel

By ELAD DEMALACH AND ANALIA SCHLOSSER

Table of Contents

Appendix A – Details on Universal Preschools

Program Structure and Staffing

Preschools operated five days per week for 32 hours total, serving up to 35 children per classroom. Each classroom was staffed by one certified teacher and one teacher aide. Teachers were required to have certification from academic institutions recognized by Israel's Ministry of Education (MOE) and were employed directly by the MOE. Teacher aides were required to have at least 12 years of education plus a teaching aide certificate and were employed by the local authorities (Ministry of Education Directive 36/2b, February 2002). Additional staffing sometimes included early childhood education students completing practical training requirements (Kimhi 2012).

Teacher Training and Preparation

Teaching education was obtained through specialized teacher training colleges that constituted the primary entry pathway to the profession. These colleges were directly supervised and financed by the MOE, with 23 colleges belonging to the state education sector. Most of these institutions offered early childhood education programs that, during the study period, focused on ages 3-8 (including first and second grade). Three colleges (located in the north district) were specifically designated for the Arab sector, while three additional state colleges maintained special tracks for Arab, Druze, and Bedouin education. Some Arab students enrolled in Hebrew-language programs at secular sector colleges (Kimhi 2012). In 1995 and 1996, prior to the law implementation, the number of training programs for Arab preschool teachers was doubled. Concurrently, the MOE increased its oversight of preschools, enhanced the quality of professional support, and introduced specialized curricula (Ministry of Justice 2001, page 293).

Infrastructure

Establishing preschools in Arab localities presented practical challenges due to limited availability of suitable physical spaces. This shortage of facilities required local authorities to explore alternative solutions and adapt to available options. To address these space limitations, they utilized public spaces owned by local municipalities, such as community centers, and supplemented these with rented buildings. When existing structures were not sufficient, they constructed additional classrooms using prefabricated buildings (Kimhi 2012). While the lack of adequate physical infrastructure posed ongoing challenges, these adaptive approaches helped facilitate preschool expansion in Arab localities.

Pedagogical Approach

The pedagogical approach of preschools followed the core program developed by the Preschool Education Division of the MOE for children aged 3-5. This program was initially translated from Hebrew and then gradually adapted to meet the specific needs and characteristics of the Arab population (see Aram and Ziv 2018 for more details). The program emphasized skill development through small-group instruction with teaching staff and whole-class learning activities, balanced with unstructured free play.

Curriculum Components

The core program included four clusters:

- **Language** - Literacy skills, development of expression and readiness for reading, writing, and comprehension.
- **Mathematics, Science and Technology** - Mathematical thinking, exposure to scientific concepts, first experiments, and familiarity with technological environments.
- **Arts** - Development of creativity and expression skills, movement, music, and plastic arts.
- **Life Skills** - Education for wellness, social skills, safety and road safety.

Program Goals

According to a position paper of the division of preschool education at the Ministry of Education (2007), preschools (ages 3-6) had the following goals:

- Narrow educational and academic gaps that tend to widen as children get older.
- Unlock each child's full potential while striving for excellence and high achievement from early childhood.
- Instill values, knowledge, and skills that ensure equitable learning opportunities and make education accessible to every child in the system.
- Develop foundational skills and knowledge that ease the kindergarten-to-school transition, ensuring educational continuity and helping children successfully integrate academically, personally, and socially into first grade.

- Create a resource-rich educational environment that provides meaningful learning experiences and opportunities for success for every child.
- Identify and detect children at social and academic risk early, providing individualized support based on their specific needs, while maintaining the pedagogical principles of kindergarten teaching
- Integrate learning, play, creativity, spontaneity, discovery-based learning, and imagination development, while preserving the joy of being in kindergarten and adapting to each child's emotional, social, and cognitive developmental stage.

Appendix B – Data

Our analysis data is based on several datasets that were merged using individual identifiers provided by the Central Bureau of Statistics (CBS). These identifiers are based on each individual's unique ID number, which is provided to all Israelis upon birth or immigration. Below is a brief description of each dataset. The exact definition of each outcome variable is summarized in Table A2.

Population registry records: include information on parent's and children's IDs, as well as basic demographic information (year and month of birth, gender, religion, country of birth, locality of residence, and marital status). These records allow us to construct indicators for parents' marital status and number of children and to allocate the child's treatment status based on birth cohort and locality of residence during childhood. Ideally, we would observe the individual's locality of residence at age 2, just before eligibility to preschool enrollment. In practice, locality of residence can be observed only in 1995 and 2000, so for some cohorts we record locality of residence at ages 3–5 (see Table A22 for the full breakdown by cohort). This could introduce some bias if parents moved to localities that offered preschool services; however, this is not a concern in our study, as migration between localities is rare among Arabs in Israel. For instance, 96% of the children in our sample born in 1991–1994 were in the same locality in 1995 (pre-reform) and 2000 (post-reform). This aligns with Hleihel (2011), who found that only 9.5% of adult Arabs in Israel lived outside their birth locality. Moreover, we do not find systematic changes in the sociodemographic characteristics of individuals in treated versus comparison localities between the pre- and post-reform periods (see Table A6 and discussion in Appendix D).

GEMS test scores: The GEMS exams (*Meitzav*) are low-stake standardized tests administered by the National Authority for Measurement and Assessment of Education (RAMA) in Israel to students in the fifth and eighth grades in four subjects: verbal skills in native language (Arabic for our sample), English, math, and science. The raw test scores use a 1-to-100 scale that we transform into z-scores to facilitate interpretation of results. Administration of the GEMS exams is designed so that only a national representative sample of schools is tested each year.[1] This design imposes some challenges for our estimation methodology. First, it implies that we have a smaller sample for the estimation of the effect of universal preschool on test scores in a given subject. Second, the cohort fixed-effect (λ_t) of our main DID specification in equation (1) is affected by the sample composition of the localities in which GEMS exams are administered for each cohort.[2] To circumvent this problem, we estimate equation (1) replacing the cohort fixed effect with a cohort-by-test-year fixed effect, effectively comparing localities that took the GEMS exams in exactly the same years.

GEMS student questionnaires: Schools participating in the GEMS exams also complete questionnaires administered to all students in grades 5–9. In these questionnaires, students are asked to indicate the extent to which they agree with a number of statements on a 6- or 5-point Likert scale ranging from 1 (strongly agree) to 5 or 6 (strongly disagree). In order to have consistent outcomes for ease of interpretation, we construct binary indicators that take a value of 1 if respondents partially to strongly agree with each statement, and 0 otherwise. Our data on student questionnaires cover the years 2002-2013. In 2007, which is roughly the middle of the sample period, the format of the student questionnaire was revised, some questions were modified, and the Likert scale was extended from 5 to 6 points. Therefore, we focus on a specific subset of questions that remained very similar or identical throughout the sample period. Note that these changes to the student questionnaire are not expected to bias our estimates for two reasons: (1) we include year fixed effects, and (2) the year of the format change does not overlap with the year of the reform implementation, as the change occurred during the pre-reform period for some cohorts and the post-reform period for others.

[1] All localities are grouped into four groups, where each group constitutes a representative sample of all Israeli schools. Each group is tested every other year in only two subjects: math and native language, or science and English (as a foreign language). Thus, students in a given school are tested in the same subject only once in four years. However, the localities in our study did not fully comply with this official test-taking calendar.

[2] As the sampling design is supposed to provide a representative sample of the entire population of schools, the potential bias should vanish for a large sample of localities that fully comply with the official test-taking calendar. However, our analysis sample encompasses a limited number of localities (37).

Matriculation exams: The data on the matriculation exams include information on all subjects that students were tested in towards their matriculation certificate in grades 10–12. The matriculation certificate is earned by passing a series of national exams in core and elective subjects. Students choose to be tested at various levels of proficiency, with each test awarding 1–5 credit units per subject, depending on difficulty. Some subjects are mandatory, and, for many, the most basic level is three credit units. Advanced level subjects are those taken at four or five credit units. A minimum of 20 credit units is required to qualify for a matriculation certificate. The matriculation certificate is a prerequisite for university admission and receiving it is one of the most economically important educational milestones. Similar high school matriculation exams are found in many countries and some states in the US. Examples include the New York Regents Examinations and the French baccalaureate exams.

Psychometric exam: The psychometric exam is a standardized test, similar to the U.S. SAT. It includes three sections: quantitative, verbal, and English and is administered in various languages, including Arabic. Admission to most higher education institutions in Israel is based on a weighted average of the matriculation average score and the psychometric exam score.

Postsecondary Education

Academic postsecondary education records: This dataset includes longitudinal records of individuals enrolled in Israeli higher education academic institutions between 1995 and 2018. Each entry corresponds to a specific year in which the individual appeared in the student registry. The dataset covers students from universities, academic colleges, and teacher training institutions. For each student, we identify their first appearance in the dataset as the year they started academic postsecondary education.

Vocational postsecondary education records: This dataset tracks individuals enrolled in *Mahat* institutions—Israel's network of public vocational and technological training colleges—between 1998 and 2018. It enables analysis of vocational education pathways outside the academic higher education system. For each student, we identify their first appearance in the dataset as the year they started vocational postsecondary education.

Juvenile criminal records: This dataset contains administrative records of criminal cases opened for youth aged 12–18 during the years 2003–2017. Each record includes the year the offense was committed and the type of offense, using a standardized statistical coding system. Offenses are grouped into broad categories, such as:

- Security and Public Order Offenses (e.g., offenses against state security [100] or public order [200])
- Offenses Against Life and Bodily Integrity (e.g., homicide [300], bodily harm [400])
- Sexual and Property Offenses (e.g., sex offenses [500], property crimes [700])
- Other Offense Categories (e.g., moral crimes [600], fraud [800], economic [900], administrative [1000], licensing [1100], miscellaneous [1200], and legal definition clauses [1300]).

For each individual we define indicators for any criminal record between ages 12 and 18 and specific indicators for the different categories.

Education registry: The Israel Education Registry is a comprehensive administrative database maintained by the CBS that tracks the educational attainment of nearly all individuals listed in the population registry, covering approximately 96% of those aged 25–69. It compiles data from various sources, including postsecondary institutions, government ministries, professional licensing bodies, and self-reported information from CBS surveys and censuses. In this study, we use the registry to construct the parental education variables for the individuals included in our sample.

Employee Income Tax Records: This dataset is an administrative file compiled by the CBS based an annual income report submitted by employers to the Israel Tax Authority. It includes comprehensive information on wage earnings and number of months of work for each individual with salaried income. In this study, we use this dataset to measure parental employment and income for the individuals included in our sample.

Appendix C - Assessing the Parallel Trends Assumption

To assess the robustness of the results to possible violations of the parallel trends assumption, we perform a sensitivity analysis suggested by Rambachan and Roth (2023). We focus on the treatment effect on the index of high school performance to summarize our results and gain statistical power. Results appear in Figure A3, where the blue line in each subfigure plots the confidence interval of the treatment effect for period 1 obtained on our DID model. Panel (a) plots the confidence intervals of the treatment effect allowing for violations of the linear pre-trend up to a parameter M (i.e., sensitivity analysis using smoothness restrictions). The figure shows that the treatment effect would still be positive and significant if we allow for the difference in trends between the treated and control groups to be linear (M=0). The breakdown value for a significant effect is at M=0.005, which is roughly 25% of the standard error of the treatment effect of the high school index. We also apply the second approach proposed by Rambachan and Roth (2023) and plot the results in Panel (b) (i.e., sensitivity analysis using relative magnitude restrictions). In this figure, we plot the confidence intervals for the treatment effect allowing for a post-treatment violation of parallel trends to be no larger than $\bar{M}$ times the maximum pre-treatment violation of the parallel trend. The breakdown point is $\bar{M} \approx 1.1$, meaning that we can rule out a null effect unless we allow for violations of parallel trends that are 1.1 times larger than the maximum violation observed in the pre-period. To sum up, both approaches suggest that our results would remain significant even if we allow for some deviations from the parallel trends assumption.

Appendix D – Details of the Robustness Checks

Inclusion of Background Characteristics and Time Trends

We first assess the sensitivity of our results to the inclusion of the set of background characteristics used in our main specification. Results are reported in Table A5. To ease comparison, main results appear in column (1). In column (2) we report estimates from a simple DID model that includes only time and locality fixed effects. Estimates from this simple specification are very similar to our baseline specification, reinforcing the assumption that the results are not driven by differential changes in observed covariates (or unobserved characteristics correlated with observed covariates) between treatment and comparison localities.

Given that the reform was implemented in localities classified with the lowest socioeconomic ranking, it could be argued that our results are driven by a convergence over time between lower and higher SES localities that could have occurred even without the opening of preschools. To assess this, we present in columns (3) and (4) of the same table estimates from a model that includes a linear time trend interacted with a locality's socioeconomic cluster (1 to 4) or socioeconomic ranking (1 to 203) (together with the baseline linear trend).[3, 4] The estimates remain

[3] The national ranking of the localities in our sample falls within the range of 8–138. The lower the ranking the lower the socioeconomic status.

[4] We do not allow for a specific linear trend for each cluster or ranking, as this would absorb most of the treatment effects (see, e.g., Meer and West, 2016; Goodman-Bacon, 2021).

largely similar to our main results. Some are smaller, but most remain significant. Note that the interaction between a time trend and socioeconomic ranking or cluster is highly correlated with the "*Exposed_preschool"* indicator, our main variable of interest, and therefore it is not surprising that some of the estimated effects are smaller.

Differential Changes in Background Characteristics

In Table A6, we examine whether children's background characteristics change differentially in treated versus comparison localities between the pre- and the post-reform period by estimating DID models that include only time and locality fixed effects using observed covariates as outcomes. All estimates are small and statistically insignificant, except for the coefficient on father's income, which shows a negative sign (suggesting a decline in income among treated cohorts in the post-reform period). This result further supports the causal interpretation of our findings. If anything, our results might be downward biased as father's income is typically positively correlated with child outcomes.

Placebo Treatment in the Pre-reform Period

We conduct a placebo analysis where we estimate the baseline DID equation on all main outcomes, including only pre-reform cohorts, and assume that the law was implemented in the middle of the pre-reform period, two years before it actually came into effect (Table A7). Most estimates are small and non-significant and have inconsistent signs across outcomes. Thus, we find no evidence of significant differential pre-reform trends between treatment and comparison localities, supporting our main identification assumption of no differential trends in the post-reform period.

Using Different Subsamples

A last check we perform relates to the experimental setup. Note that our comparison group is composed of two different groups of localities: those that did not receive universal preschool education during the period we cover in this study (never treated) and those that already had preschool education before implementation of the law due to their special status (always treated). In some settings, such as a staggered DID design, it is problematic to use early-treated units as a comparison group for late-treated units (e.g., Callaway and Sant'Anna, 2021; Roth et al., 2023). We explain in Section 2 in the main text why this is less of a concern in our setup. Nevertheless, we report in Table A8 the results of the estimation where we use only one specific group of localities as a comparison group: never treated (column (2)) or always treated (column (3)). To ease comparison, our main estimates appear in column (1). Overall, most of our results hold when we use only one type of localities as a comparison group.

In columns (4) to (6) of the same table, we assess the robustness of our results with respect to additional issues related to sample composition. As we have a relatively small sample of localities (37), we want to ensure that our results do not derive from a particular group of localities. We first re-estimate our model by omitting the city of Nazareth, which accounts for 16% of the sample, and is by far the largest Arab locality in the sample (column (4)). We then re-estimate our model omitting all Druze localities, all of which are in the comparison group (column (5)). Finally, we re-estimate our model omitting all Bedouin localities, most of which are in the treatment group (column (6)). Despite these changes in the composition of the localities in our sample, all estimates are highly similar to our main results, providing further support for the validity of our identification strategy. The robustness of our results across these different subsamples also suggests that our findings are not driven by ethnic-specific trends within the Arab community in Israel.

As an additional check to assess the sensitivity of our results, we re-estimate our model by dropping one locality each time to ensure that our main results do not derive from any particular locality. In Figure A6 we plot estimates along 95% confidence intervals for our main outcomes from these subsamples along with our main results. Taken as a whole, all figures indicate that our main results do not derive from any particular locality.

Clustered Standard Errors

Given our relatively small number of clusters (37 total, 15 treated), we address potential small sample bias in clustering standard errors by implementing the wild bootstrap procedure (Cameron et al., 2008). Table A9 shows that p-values obtained using standard cluster adjustments and those from wild bootstrap are very similar across outcomes. These results confirm that our statistical inference is robust to the clustering method employed.

Changes in Other School Inputs or Resources

An additional concern is that other changes might have taken place during the study period that could have affected the performance of children in treatment or comparison localities. In particular, we are concerned about other differential investments in educational inputs across treatment and comparison localities. We examine one such potential input: average class size. Using supplemental data from local authorities' statistical yearbooks compiled by the CBS, we compute average class size for individuals in both the pre- and post-reform cohorts throughout their elementary, middle, and high school years and estimate a simple DID specification that includes locality and cohort fixed effects using average class size as an outcome. Estimates for the post-reform cohorts in treatment localities, reported in Table A10, are inconsistent across schooling stages and none of them are statistically or economically significant.

We also examine whether there were other differential changes in resources in treatment versus comparison localities. For this purpose, we compiled additional data from local authorities' statistical yearbooks compiled by the CBS to examine potential differential changes in per capita expenditure, expenditure in education per capita (ages 0-17), and revenue per capita between treatment and comparison localities.[5] Although the earliest available data is from 1999, precluding analysis of pre-trends, we can assess whether these variables increased differentially between 1999 and later years. Table A11 presents DID estimates for these variables obtained from a simple model that includes year and locality fixed effects and the interaction between treatment and an indicator for the post-reform years (2000 onwards). Overall, there is no evidence of differential increases in per capita expenditure or revenue in treated localities after 1999.

However, even if there are no differential changes in class size or other local investments between treated and comparison localities that overlap with the provision of universal preschool, a second concern could arise if resources increased, as long as they had larger effects in more disadvantaged students, given that treated localities are poorer. Indeed, during the period examined, class size declined by a similar magnitude in both treatment and comparison localities (Figure A7), while expenditure and revenue per capita increased modestly (Figure A8). Nevertheless, these changes cannot explain our results given that while the decline in class size occurred gradually over time, the event study figures show a sudden, discontinuous increase in outcomes for the cohorts exposed to universal preschool. Moreover, for other investments in treated localities to bias our results, they would need to differentially affect children aged four or younger relative to children aged 5-9—an unlikely scenario.[6]

[5] We use ages 0-17 to normalize expenditure in education as these are the official population counts reported by the CBS.

[6] Note also that our placebo analysis finds no significant effects when we estimate a DID model using only pre-reform cohorts and assume the law was implemented mid-period (Table A7). If our results were driven by differential effects of additional school inputs affecting poorer areas more strongly, we would expect to find spurious treatment effects in this falsification test. Finally, we continue to find significant effects of universal preschool when we focus exclusively on the most disadvantaged children from both treated and comparison localities, whether identified by background characteristics or predicted outcomes (Tables 6 and 7), who presumably would be similarly affected by any additional school inputs.

Late-Treated Localities

As noted in the background section (Section 1), some localities were added in subsequent years due to a change in their socioeconomic cluster (i.e., they were reclassified into clusters 1 and 2): two localities were included in 2001 and three in 2003. We excluded these five localities from our main analysis sample because we do not observe their outcomes beyond high school. Moreover, we have fewer treated cohorts for which to measure their outcomes (e.g., only one cohort for those treated in 2003). Nevertheless, we perform here a secondary analysis to assess the robustness of our results when these five localities are included. In Table A12 we report our main results for high school outcomes. Column (1) displays estimates from our main sample and column (2) shows estimates after adding these localities.[7] Overall, the two sets of estimates are highly similar, confirming the causal interpretation of our findings and minimizing concerns that our results are confounded by a specific shock that affected the treated cohorts in 1999.

Appendix E – Analysis of Selection into Identification in Family Fixed Effects Model

As noted by Miller et al. (2023), the family fixed effects model identifies impacts for "switcher" families (those with children of preschool age in both the pre- and post-reform periods). These families may differ from the broader population affected by universal preschool, potentially affecting treatment effect estimates. We address this point by comparing three groups: our main sample, the sibling sample (i.e., individuals with at least one sibling in the sample), and the "switcher" sample, focusing on pre-treatment cohorts (Table A13). In our study, 54% of the children come from "switcher" families—substantially higher than the 4% reported by Miller et al. (2023) for Head Start families. These children are from slightly more disadvantaged backgrounds with somewhat worse outcomes, though differences are minor. For example, average family size for "switchers" is 3.32 compared to 3.14 in the sibling sample, and 3.07 in the main sample. Average maternal education is 10.04 in the "switcher" sample, compared to 10.24 in the sibling sample, and 10.28 in the main sample. Almost half (48%) of the children from "switcher" families earned a matriculation certificate versus 49% in the sibling sample and 50% in the main

[7] The estimating equation is identical to equation (1), where the $Exposed\_Preschool_{s(t+4)}$ indicator gets the value of 1 for the relevant exposed cohorts in these five additional localities.

sample. Given these relatively small differences in background characteristics between the "switchers" sample and the main sample, we do not expect estimates from the family fixed effects model to be affected by sample composition. Indeed, our main results remain consistent in this subsample.

Appendix F – Impact on Maternal Employment and Earnings

We examine the impact of universal preschool on maternal employment and earnings using two approaches. We first estimate the same DID model (equation (1)) based on our main children's sample, using as outcomes several measures of mothers' labor market outcomes: indicators for mother's employment at ages 3–5, number of months worked, and log wages. In addition to the main controls, the model also controls for mother's age and age squared.

Results appear in Table A18, with estimates for the full sample in column (1) and estimates for subsamples stratified by mothers' education in columns (2) and (3). The employment rate of mothers of children aged 3–5in the pre-reform period was extremely low: 17%. The employment rate of mothers with less than a high school education (who account for 60% of our sample) is even lower: 11%. Overall, there was no change in employment rates, months worked, or wages among mothers of children who received universal preschool. Estimates for all outcomes are positive but small and are not statistically significant.

As an alternative strategy, we use the mothers as a unit of analysis and estimate DID models comparing labor market outcomes of mothers of children aged 3-5 five years before and after implementation of universal preschool in treated and comparison localities (1995–2004).[8] Such a strategy allows us to compare the effects of preschool exposure among mothers of preschool-aged children with a "placebo" effect among other mothers of children who are not preschool-aged in the same set of localities. We can thus rule out the possibility that results are spuriously driven by time-varying labor market conditions that differentially affected treatment and comparison localities, such as the 2001-2002 recession in Israel.[9] As in the previous analysis, we estimate the models using the full sample and subsamples stratified by mothers' education (Table A19).

[8] We also select mothers of children aged 5 in this sample, as most children turn 5 while attending preschool (the cutoff date for entering grade 1 was around September 1).

[9] Israel experienced an economic recession in these years due to two main factors: (1) outburst of the violent period of the second Intifada (2) bust of the Dot-com bubble (see Bank of Israel, 2002, 2003).

Estimates show no significant effects of universal preschool provision on the labor supply or wages of mothers of children aged 3–5, nor for mothers who have children of other ages. We therefore conclude that universal preschool had no significant effect on mothers' employment or income during the period analyzed in this study. As a result, we can rule out increases in mothers' employment and income as possible channels that could explain the positive impacts we find on children's outcomes.

REFERENCES

Aram, D. and Ziv, M. Early Childhood Education in Israel (2018). In J. L. Roopnarine., J. E. Johnson., S. Quinn., & M. Patte (Eds.). International Handbook of Early Childhood Education. New York: Routledge. Chap. 8.

Bank of Israel. 2002. Recent Economic Developments, 99, April-September 2002. Bank of Israel.

Bank of Israel. 2003. Recent Economic Developments, 100, July-December 2002. Bank of Israel.

Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller. 2008. "Bootstrap-Based Improvements for Inference with Clustered Errors." *Review of Economics and Statistics* 90 (3): 414–27.

García, Jorge Luis, James J. Heckman, and Victor Ronda. 2023. "The Lasting Effects of Early-Childhood Education on Promoting the Skills and Social Mobility of Disadvantaged African Americans and their Children." *Journal of Political Economy* 131 (6): 1477–506.

García, Jorge Luis, James J. Heckman, and Anna L. Ziff. 2018. "Gender Differences in the Benefits of an Influential Early Childhood Program." *European Economic Review* 109: 9–22.

Goodman-Bacon, Andrew. 2021. "Difference-in-Differences with Variation in Treatment Timing." *Journal of Econometrics* 225 (2): 254–77.

Israeli Ministry of Justice (2001). *Initial periodic report to the Committee on the Rights of the Child* (CRC/C/8/Add.44). Submitted to the United Nations Committee on the Rights of the Child.

Israeli Ministry of Education (2007). [in Hebrew]. "Implementation Program of the early childhood Education Division", Position Paper.

Meer, Jonathan, and Jeremy West. 2016. "Effects of the Minimum Wage on Employment Dynamics." *Journal of Human Resources* 51 (2): 500–22.

Kimhi, A. (Ed.). (2012). Pre-primary education in Israel: Organizational and demographic perspectives. Jerusalem: Taub Center for Social Policy Studies in Israel.

Appendix Figures and Tables

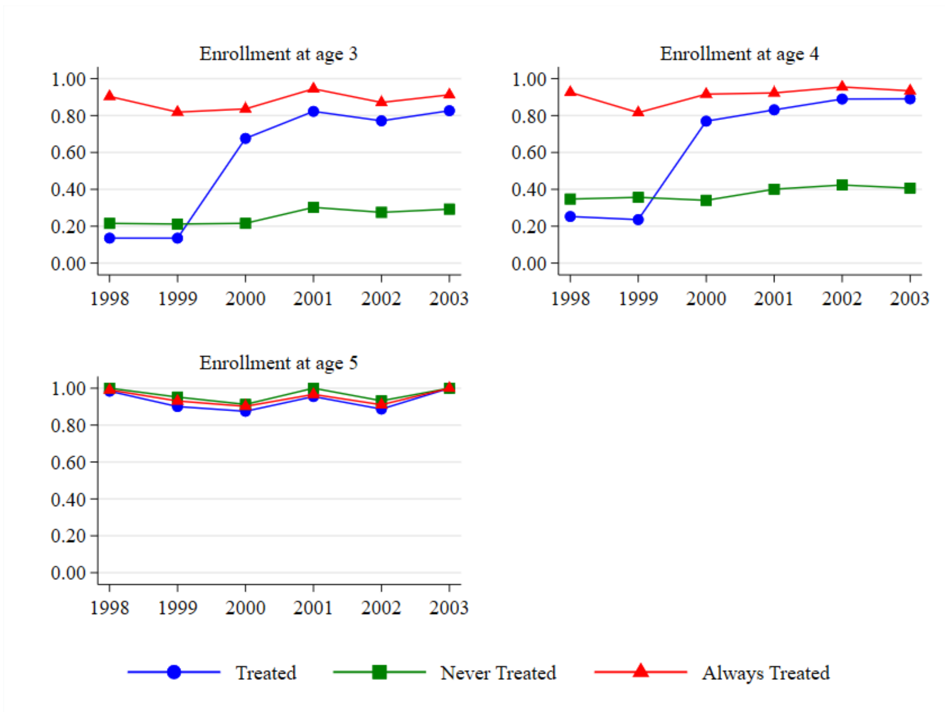


FIGURE A1. PRESCHOOL ENROLLMENT IN ARAB LOCALITIES IN ISRAEL - 1998-2003

Notes: The figure shows preschool enrollment rates of Arab children by year in different groups of localities, according to their treatment status. The analysis is based on aggregated enrollment and population counts data by locality and year provided by the Israeli Central Bureau of Statistics. Treated localities received universal preschool education starting from the year 2000. Never-treated localities are those that were not included in the first phase of the Law implementation. Always Treated localities include localities that received preschool subsidies before the Law implementation.

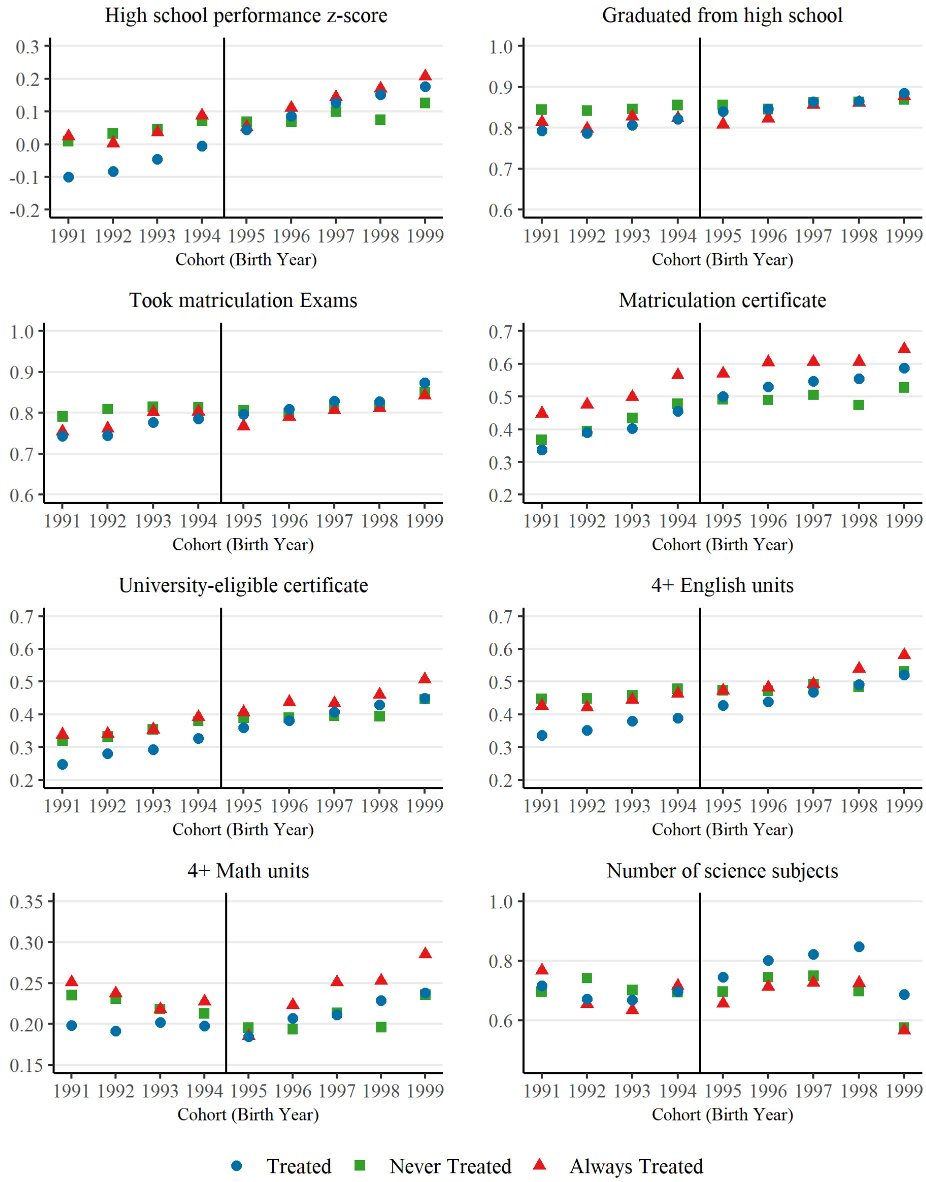
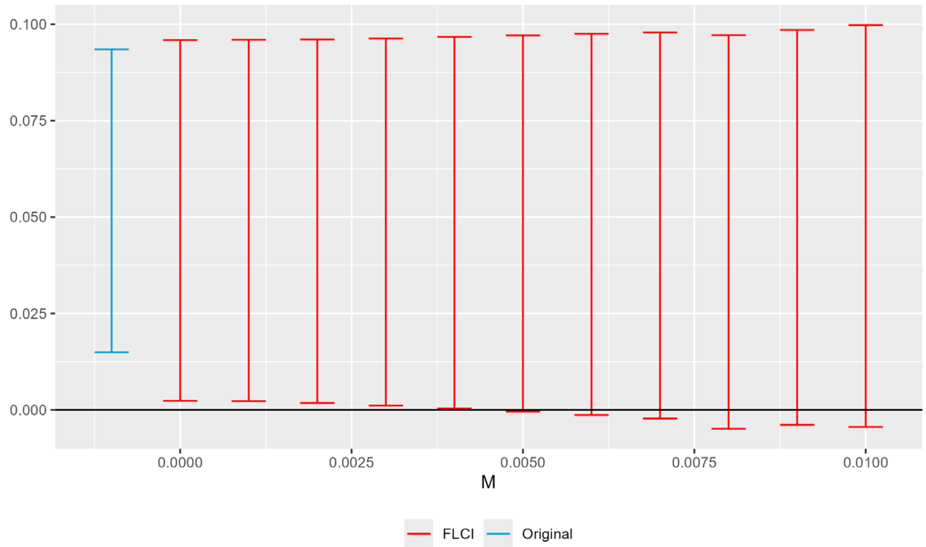


FIGURE A2. UNCONDITIONAL COHORT MEANS, BY TREATMENT STATUS

Notes: The figure shows unconditional cohort means of high school outcomes according to the locality treatment status. Treated localities received universal preschool education starting from the year 2000 (1995 cohort). Never-treated localities are those that were not included in the first phase of the Law implementation. Always-treated localities include localities that received preschool subsidies before the Law implementation.

(a) Smoothness Restriction



(b) Bounds on Relative Magnitude

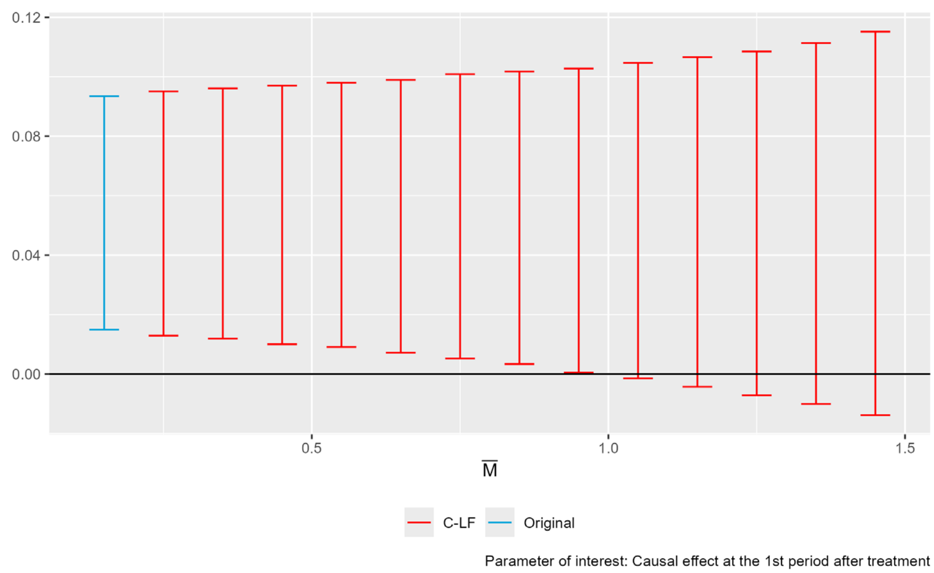


FIGURE A3. SENSITIVITY ANALYSIS FOR THE TREATMENT EFFECT ON HIGH SCHOOL PERFORMANCE TO VIOLATIONS OF THE PARALLEL TRENDS ASSUMPTION

Notes: The figure reports 90% confidence intervals for the effect of universal preschool on the index of high school outcomes in the first period after the reform in Blue and a sensitivity analysis for the effect under possible violations of the parallel trends assumptions following the method proposed by Rambachan and Roth (2023). Panel (a) plots in red the confidence intervals of the treatment effect allowing for violations of the linear pre-trend up to a parameter M (sensitivity analysis using smoothness restrictions). Panel (b) plots the confidence intervals for the treatment effect allowing for a post-treatment violation of parallel trends to be no larger than $\overline{M}$ times the maximum pre-treatment violation of the parallel trend (sensitivity analysis using relative magnitudes restrictions).

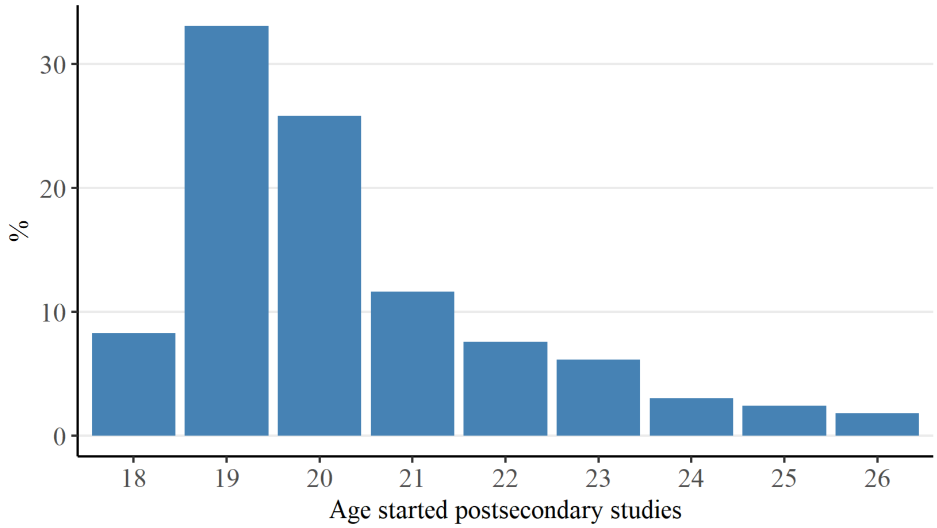


FIGURE A4. AGE DISTRIBUTION AT ENROLLMENT IN POSTSECONDARY INSTITUTIONS

Notes: The figure reports the age distribution at first enrollment in a postsecondary education institution for the prereform cohort (born in 1991) in the localities of this study. Enrollment data is available until the 2017-2018 academic year.

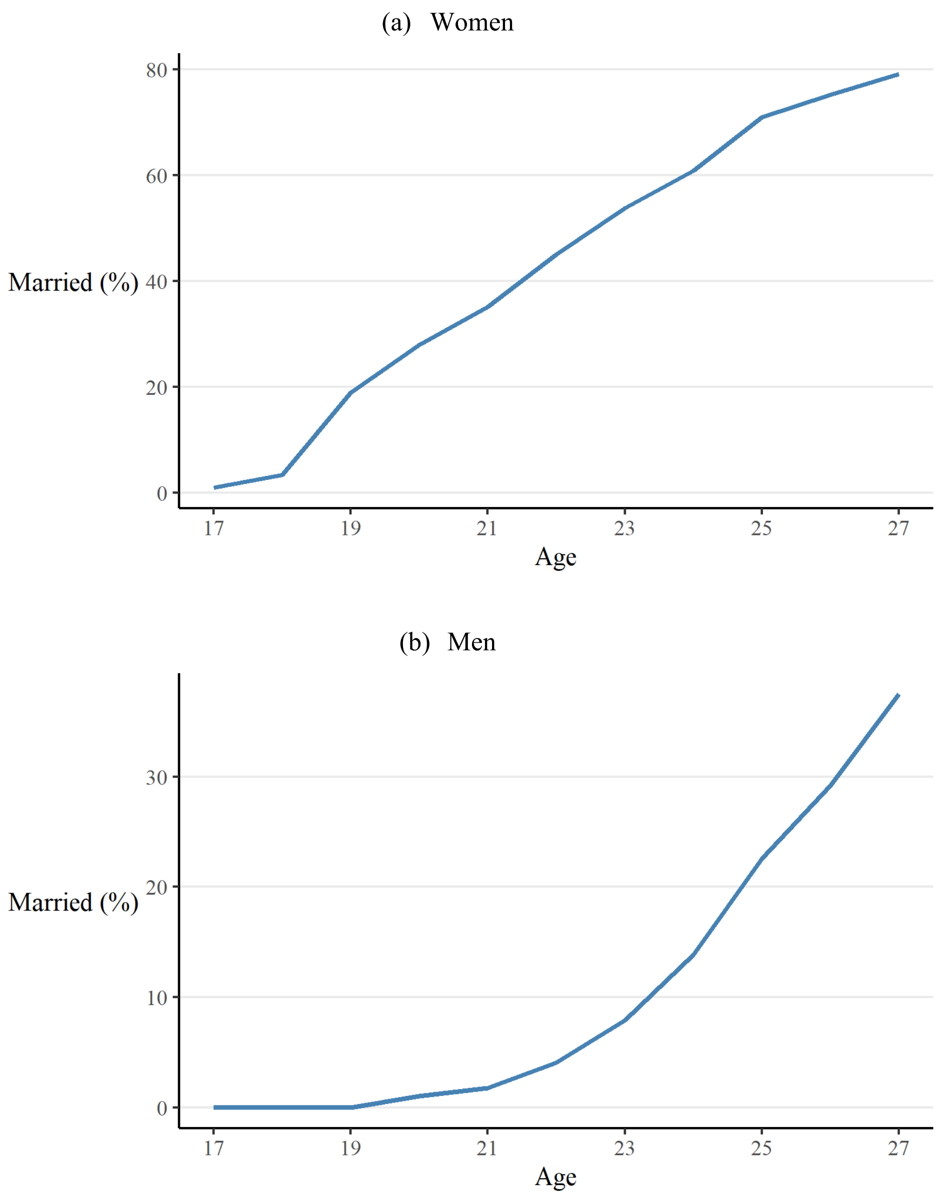


FIGURE A5. SHARE OF MARRIED INDIVIDUALS, BY AGE

Notes: The figure plots the share of married individuals by age for the prereform cohort born (born in 1991) in the localities of this study.

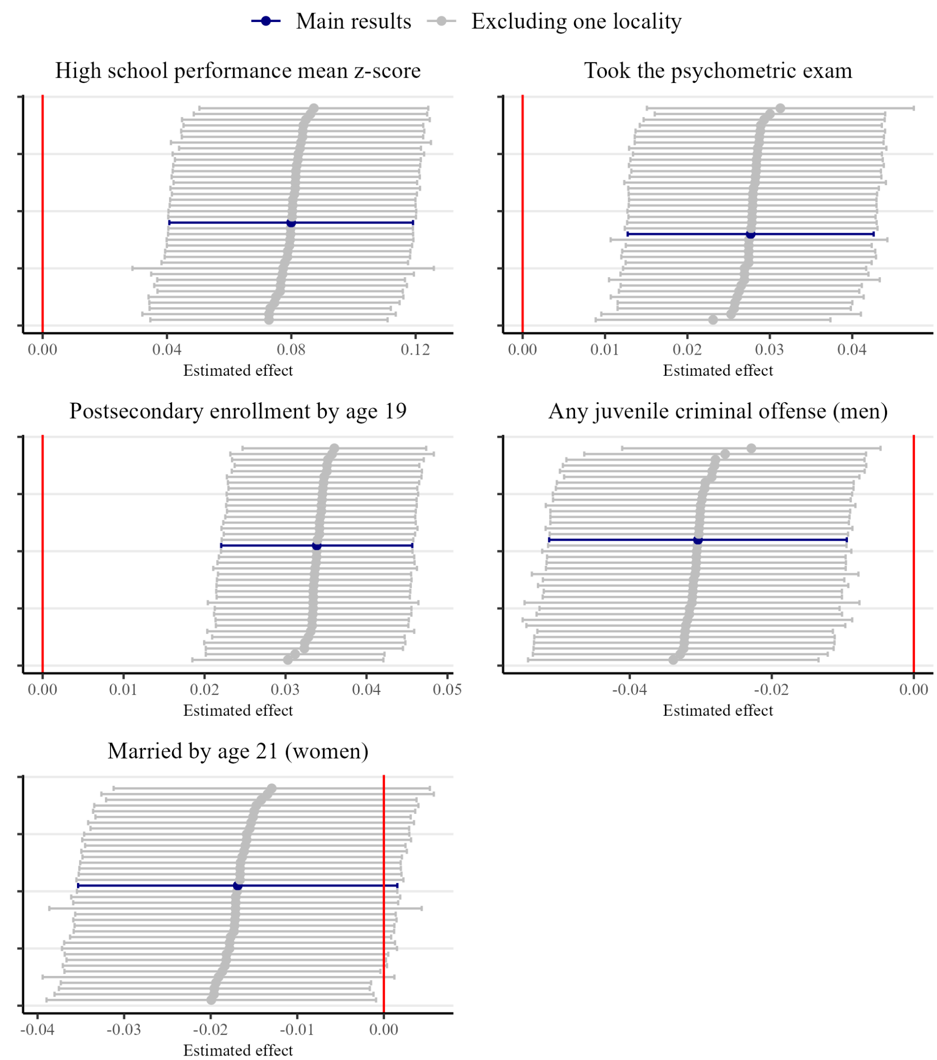


FIGURE A6. SENSITIVITY ANALYSIS OF THE IMPACT OF UNIVERSAL PRESCHOOL

Notes: The figures plot the distribution of estimates and 95\% confidence intervals of our baseline DID specification in equation (1). The blue bars represent estimates for our main sample, and the grey bars represent estimates obtained by excluding one locality from the sample at a time. The specification includes locality and cohort fixed effects and controls for parental education, mother's employment and father's earnings (in deciles) when the child was 2 years old, number of siblings, and religion. The confidence intervals are constructed with standard errors clustered at the locality level.

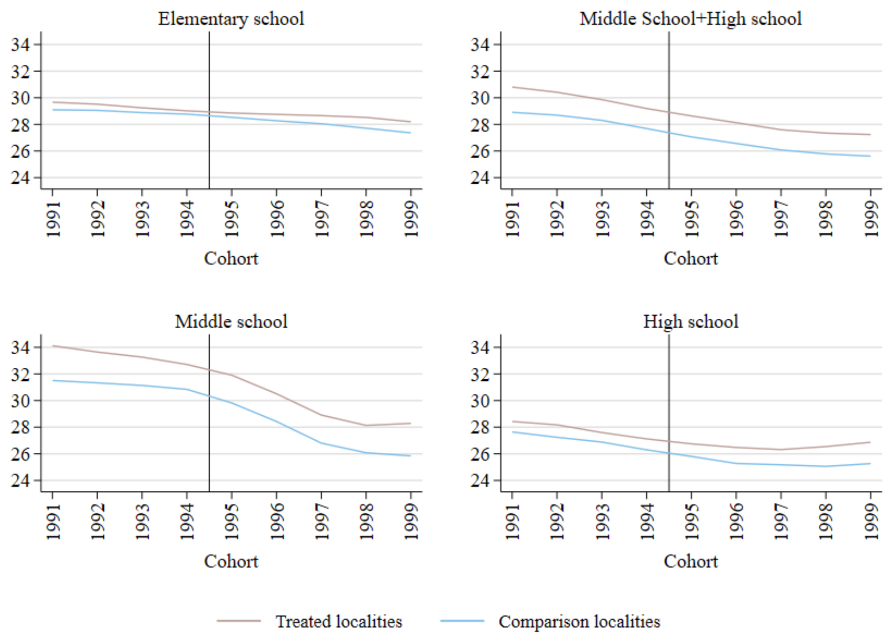


FIGURE A7. AVERAGE CLASS SIZE IN TREATED AND COMPARISON LOCALITIES

Notes: The figures plot the average class size in treated and comparison localities for the cohorts included in our sample. The data was compiled from the official reports of the Israeli Central Bureau of Statistics on the local authorities.

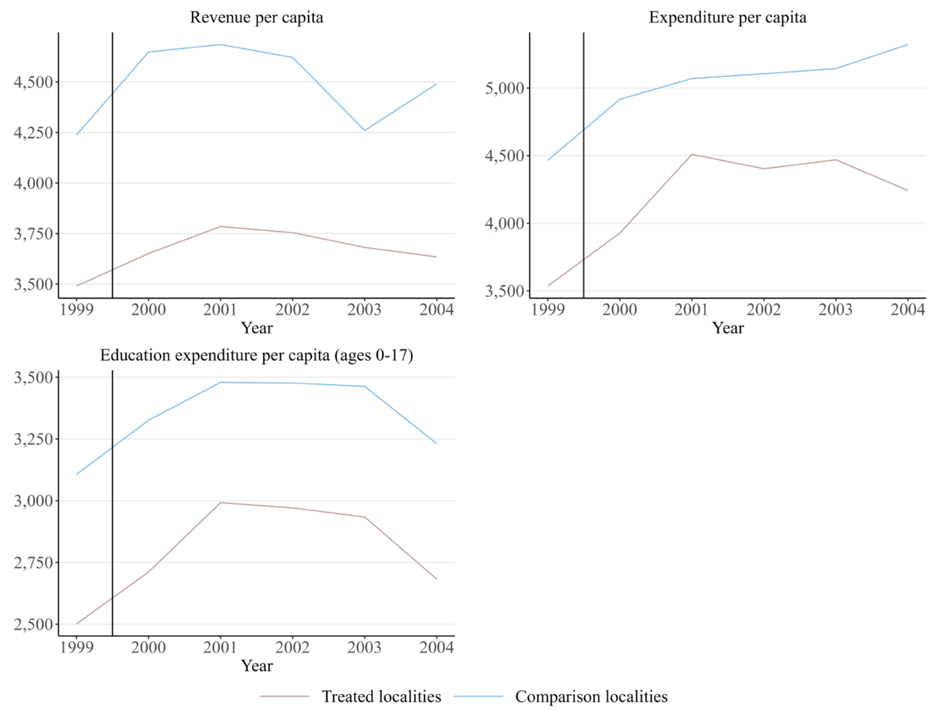


FIGURE A8. REVENUE AND EXPENDITURE PER CAPITA IN TREATED AND COMPARISON LOCALITIES

Notes: The figures plot average revenue and expenditure per capita, and average expenditure on education per capita (ages 0-17) in treated and comparison localities. The data was compiled from the official reports of the Israeli Central Bureau of Statistics on the local authorities for the years 1999-2004. The variables are reported in nominal terms in NIS. One treated and one comparison locality lack financial data in the official reports.

Table A1: Pre-reform and Post-reform Cohorts of the Study by Age

Birth Cohort									Age	Outcomes	
Pre-reform cohorts				Post-reform cohorts							
1991	1992	1993	1994	1995	1996	1997	1998	1999			
1993	1994	1995	1996	1997	1998	1999	2000	2001	1-2		
1994	1995	1996	1997	1998	1999	2000	2001	2002	2-3		
1995	1996	1997	1998	1999	2000	2001	2002	2003	3-4		
1996	1997	1998	1999	2000	2001	2002	2003	2004	4-5		
1997	1998	1999	2000	2001	2002	2003	2004	2005	5-6		
1998	1999	2000	2001	2002	2003	2004	2005	2006	6-7		
1999	2000	2001	2002	2003	2004	2005	2006	2007	7-8		
2000	2001	2002	2003	2004	2005	2006	2007	2008	8-9		
2001	2002	2003	2004	2005	2006	2007	2008	2009	9-10		
2002	2003	2004	2005	2006	2007	2008	2009	2010	10-11	GEMS 5	
2003	2004	2005	2006	2007	2008	2009	2010	2011	11-12		
2004	2005	2006	2007	2008	2009	2010	2011	2012	12-13		
2005	2006	2007	2008	2009	2010	2011	2012	2013	13-14	GEMS 8	
2006	2007	2008	2009	2010	2011	2012	2013	2014	14-15		Juvenile crime
2007	2008	2009	2010	2011	2012	2013	2014	2015	15-16		
2008	2009	2010	2011	2012	2013	2014	2015	2016	16-17		
2009	2010	2011	2012	2013	2014	2015	2016	2017	17-18		
2010	2011	2012	2013	2014	2015	2016	2017	2018	18-19		
2011	2012	2013	2014	2015	2016	2017	2018		19-20		
2012	2013	2014	2015	2016	2017	2018			20-21	High school graduation, matriculation, psychometric exams, Postsecondary enrollment, Marriage	
2013	2014	2015	2016	2017	2018				21-22		
2014	2015	2016	2017	2018					22-23		
2015	2016	2017	2018						23-24		
2016	2017	2018							24-25		
2017	2018								25-26		
2018									26-27		

Note: This table shows the pre-reform and postreform cohorts of the study and their ages at different years in which the outcomes of the study are measured.

Table A2: Description of the Outcome Variables

Variable name	Variable description
High School	
Graduated from high school	=1 if individual was enrolled in 12^{th} grade; 0 otherwise
Took matriculation exams	=1 if individual took at least one matriculation exam; 0 otherwise
Matriculation certificate	=1 if individual earned a Matriculation certificate; 0 otherwise
University-eligible certificate	=1 if individual earned a Matriculation certificate with at least 3 units in math and 4 units in English; 0 otherwise
4+ English units	Four or more matriculation units earned in English (0-5).
4+ math units	Four or more matriculation units earned in math (0-5).
Number of science subjects	Number of science subjects taken, as defined by the Israel Ministry of Education: physics, chemistry, biology, and computer science.
Psychometric Exam	
Took the psychometric exam (any time/by age 19)	=1 if individual took the psychometric exam at least once; 0 otherwise (any time/ by age 19)
Psychometric total score indicators	Indicators for obtaining a total score at or above the 1st, 2nd, or 3rd quartile (400, 470, 580)
Psychometric verbal score indicators	Indicators for obtaining a score in the verbal section (Arabic) at or above the 1st, 2nd, or 3rd quartile (80, 93, 109)
Psychometric quantitative score indicators	Indicators for obtaining a score in the quantitative section at or above the 1st, 2nd, or 3rd quartile (85, 99, 119)
Psychometric English score indicators	Indicators for obtaining a score in the English section at or above the 1st, 2nd, or 3rd quartile (78, 88, 107)
Postsecondary Outcomes	
Postsecondary enrollment	=1 if individual was enrolled in any Israeli postsecondary institution; 0 otherwise
Academic institution	=1 if individual was enrolled in any postsecondary institution with academic degree credentials (university, academic college, or teacher training institution) ; 0 otherwise
University (first tier)	=1 if individual was enrolled in a university, which is a first-tier academic institution in Israel; 0 otherwise
Academic college	=1 if individual was enrolled in an academic college, which is a second-tier academic institution in Israel; 0 otherwise
Teacher training institution	=1 if individual was enrolled in a teacher training institution; 0 otherwise
Vocational institution	=1 if individual was enrolled in a postsecondary vocational or technological training college; 0 otherwise
Juvenile Crime	
Any juvenile criminal offense	=1 if individual had at least one criminal offense by age 18; 0 otherwise
Security/order criminal offense	=1 if individual had at least one criminal security or order offense by age 18; 0 otherwise
Life/body criminal offense	=1 if individual had at least one criminal life or body offense by age 18; 0 otherwise
Sex/property criminal offense	=1 if individual had at least one criminal sex or property offense by age 18; 0 otherwise
Other criminal offense	=1 if individual had at least one criminal offense in other categories by age 18; 0 otherwise
Marriage	
Married by age 18/19/20/21	=1 if individual was officially married according to the Israel Marriage Registry by age 18, 19, 20, or 21
GEMS exam (*Meitzav*)	
Arabic (native) language grade	Grade in the Arabic language GEMS exam (in terms of s.d. units, original scale is 0-100)
Math grade	Grade in the math GEMS exam (in terms of s.d. units, original scale is 0-100)
English grade	Grade in the English GEMS exam (in terms of s.d. units, original scale is 0-100)
Science grade	Grade in the science GEMS exam (in terms of s.d. units, original scale is 0-100)

Table A3: Descriptive Statistics - Treatment and Comparison Localities

	Treatment (1)	Comparison (2)	Difference (3)
Population size	8,865	9,564	-700
	(6,090)	(12,550)	(3,109)
Median age	18.33	21.90	-3.57
	(1.50)	(2.59)	(0.70)
Dependency ratio	121.69	102.79	18.9
	(14.71)	(12.74)	(4.74)
Families with 4 or more children (%)	0.40	0.30	0.1
	(0.08)	(0.09)	(0.03)
Income per capita	1,237	1,465	-228
	(125)	(374)	(90)
Rate of motorization	0.14	0.18	-0.04
	(0.02)	(0.04)	(0.01)
New motor vehicles (%)	0.16	0.18	-0.02
	(0.04)	(0.04)	(0.01)
Students among aged 20-29 (%)	0.04	0.08	-0.05
	(0.02)	(0.04)	(0.01)
Entitled to matriculation certificate among aged 17-18 (%)	0.28	0.42	-0.14
	(0.09)	(0.16)	(0.04)
Earners below minimum wage (%)	0.55	0.51	0.03
	(0.04)	(0.06)	(0.02)
Earners above twice average wage (%)	0.01	0.03	-0.01
	(0.00)	(0.01)	(0.00)
Recipients of income support (%)	0.03	0.02	0.01
	(0.01)	(0.01)	(0.00)
Recipients of income supplements to old age pension (%)	0.46	0.27	0.19
	(0.09)	(0.07)	(0.03)
Number of localities	15	22	

Notes: This table presents descriptive statistics and balancing tests between the treatment and comparison localities based on characteristics from 1999. Columns (1) and (2) display the means (and standard deviations (in parentheses)) in each category. The differences in means between treatment and comparison localities appear in Column (3), with robust standard errors (in parentheses).

Table A4: Descriptive Statistics pre-reform Cohorts

	Treatment	Comparison	Difference		Treatment	Comparison	Difference
	(1)	(2)	(3)		(1)	(2)	(3)
Panel A: pre-treatment covariates				**Panel B: outcomes**			
Father's years of education	9.92	10.65	-0.73	Completed high school	0.80	0.83	-0.03
	(3.19)	(3.20)	(0.24)		(0.40)	(0.37)	(0.03)
Mother's years of education	9.42	10.13	-0.71	Participated in the matriculation exams	0.76	0.80	-0.03
	(3.09)	(3.04)	(0.38)		(0.43)	(0.40)	(0.03)
Father employed in 1998	0.67	0.66	0.01	Matriculation certificate	0.40	0.46	-0.06
	(0.47)	(0.47)	(0.02)		(0.49)	(0.50)	(0.04)
Mother employed in 1998	0.13	0.18	-0.05	University-eligible matric. certif.	0.29	0.35	-0.06
	(0.33)	(0.38)	(0.02)		(0.45)	(0.48)	(0.02)
Father's monthly wages in 1998	4,942	5,942	-1,001	4+ English units	0.36	0.45	-0.09
	(3,926)	(4,781)	(177)		(0.48)	(0.50)	(0.03)
Mother's monthly wages in 1998	2,741	2,973	-232	4+ math units	0.20	0.23	-0.03
	(1,976)	(2,368)	(163)		(0.40)	(0.42)	(0.02)
Number of siblings	3.65	3.06	0.59	Number of science subjects	0.51	0.52	-0.01
	(2.11)	(1.80)	(0.14)		(0.74)	(0.70)	(0.07)
Share of females	0.49	0.48	0.00	Any juvenile criminal record (men)	0.17	0.13	0.03
	(0.50)	(0.50)	(0.00)		(0.37)	(0.34)	(0.02)
Share of Druze	0.00	0.25	-0.25	Took the psychometric exam	0.39	0.41	-0.02
	(0.01)	(0.43)	(0.09)		(0.49)	(0.49)	(0.03)
Share of Bedouin	0.21	0.03	0.18	Average psychometric score	472	484	-12
	(0.40)	(0.17)	(0.10)		(112)	(113)	(8)
				Any postsecondary enrollment	0.33	0.39	-0.06
					(0.47)	(0.49)	(0.03)
Number of localities	15	22		Married by age 21 (women)	0.32	0.22	0.09
Number of observations	14,442	21,226			(0.47)	(0.42)	(0.04)

Notes: This table presents descriptive statistics and balancing tests between treatment and comparison groups for various characteristics of the pre-reform cohorts. Columns (1) and (2) display the means (and standard deviation (in parentheses)) in each category. The differences in means between the treatment and comparison localities are reported in Column (3), with standard errors clustered at the locality level.

Table A5: Robustness Checks - Alternative Specifications

Dependent Variable	Main results (1)	No controls (2)	Linear trends X SES ranking (3)	Linear trends X SES cluster (4)
High school performance z-score	0.08	0.098	0.065	0.075
	(0.020)	(0.024)	(0.023)	(0.026)
	-0.058	*-0.058*	*-0.058*	*-0.058*
Took the psychometric exam	0.028	0.037	0.019	0.022
	(0.008)	(0.009)	(0.008)	(0.008)
	0.389	*0.389*	*0.389*	*0.389*
Postsecondary enrollment by age 19	0.034	0.037	0.028	0.028
	(0.006)	(0.007)	(0.007)	(0.009)
	0.157	*0.157*	*0.157*	*0.157*
Any juvenile criminal offense (men)	-0.03	-0.033	-0.036	-0.033
	(0.011)	(0.011)	(0.013)	(0.013)
	0.165	*0.165*	*0.165*	*0.165*
Married by age 21 (women)	-0.017	-0.021	0.004	0.003
	(0.009)	(0.010)	(0.011)	(0.011)
	0.318	*0.318*	*0.318*	*0.318*
Number of localities	37	37	37	37
Number of observations	84,425	84,425	84,425	84,425

Notes: This table shows various robustness checks. Column (1) reproduces our main results. Column (2) reports estimates from a simple DID specification, controlling only for locality and cohort fixed effects. Columns (3) and (4) report estimates from our main specification that controls also for an interaction between the socioeconomic ranking/cluster of the locality and a time trend. Mean outcomes of the pre-reform cohorts (born in 1991-1994) in the treatment localities appear in italics. Standard errors (in parentheses) are clustered at the locality level.

Table A6: DID Estimates to Assess Differential Changes in Background Characteristics

Dependent Variable	
Female	0.000
	(0.006)
	0.485
Druze	0.003
	(0.002)
	0.148
Bedouin	0.001
	(0.002)
	0.101
Number of siblings	-0.149
	(0.162)
	3.296
Mother Employed at age 2	0.012
	(0.010)
	0.191
Father's income above median at age 2	-0.027
	(0.012)
	0.586
Father's years of education>12	-0.003
	(0.006)
	0.162
Mother's years of education>12	0.005
	(0.006)
	0.104
Number of localities	37
Number of observations	84,425

Notes: This table shows DID estimates of the effect of universal preschool on individuals' background characteristics. The specification includes a post × treatment interaction and locality and cohort fixed effects. Mean outcomes of the pre-reform cohorts (born in 1991-1994) in the treatment localities appear in italics. Standard errors (in parentheses) are clustered at the locality level.

Table A7: Robustness Checks - Placebo Timing of Treatment

Dependent Variable	Main results (1)	Pre-reform "placebo" effect (2)
High school performance z-score	0.08	0.001
	(0.020)	(0.016)
	-0.058	*-0.091*
Took the psychometric exam	0.028	0.016
	(0.008)	(0.011)
	0.389	*0.378*
Postsecondary enrollment by age 19	0.034	0.015
	(0.006)	(0.008)
	0.157	*0.145*
Any juvenile criminal offense (men)	-0.03	0.011
	(0.011)	(0.012)
	0.165	*0.167*
Married by age 21 (women)	-0.017	-0.010
	(0.009)	(0.013)
	0.318	*0.348*
Number of localities	37	37
Number of observations	84,425	35,668

Notes: This table shows our main results for selected outcomes (column 1) and estimates of the placebo effect of universal preschool (column 2). The sample for the placebo treatment includes only pre-reform cohorts. The placebo treatment is defined for 1998 - two years before actual treatment. The specification includes locality and cohort fixed effects, controlling for parental education, mother's employment and father's earnings (in deciles) at age 2, number of siblings and religion. Mean outcomes of the pre-reform cohorts (born in 1991-1994) in the treatment localities appear in italics. Standard errors (in parentheses) are clustered at the locality level.

Table A8: Robustness Checks - Alternative Comparison Groups

Dependent Variable	Main Sample (1)	Never Treated (2)	Always Treated (3)	No Nazareth (4)	No Druze (5)	No Bedouin (6)
High school performance z-score	0.08	0.094	0.061	0.077	0.088	0.087
	(0.020)	(0.017)	(0.030)	(0.025)	(0.019)	(0.021)
	-0.058	*-0.043*	*-0.039*	*-0.050*	*-0.040*	*-0.057*
Took the psychometric exam	0.028	0.02	0.037	0.031	0.024	0.035
	(0.008)	(0.006)	(0.011)	(0.008)	(0.007)	(0.007)
	0.389	*0.389*	*0.389*	*0.389*	*0.389*	*0.403*
Postsecondary enrollment by age 19	0.034	0.035	0.031	0.03	0.031	0.036
	(0.006)	(0.007)	(0.007)	(0.006)	(0.007)	(0.007)
	0.157	*0.157*	*0.157*	*0.157*	*0.157*	*0.174*
Any juvenile criminal offense (men)	-0.03	-0.023	-0.04	-0.032	-0.023	-0.032
	(0.011)	(0.010)	(0.013)	(0.012)	(0.010)	(0.012)
	0.165	*0.165*	*0.165*	*0.165*	*0.165*	*0.161*
Married by age 21 (women)	-0.017	-0.017	-0.017	-0.017	-0.022	-0.02
	(0.009)	(0.008)	(0.014)	(0.011)	(0.009)	(0.010)
	0.318	*0.318*	*0.318*	*0.318*	*0.318*	*0.310*
Number of localities	37	20	32	36	29	30
Number of observations	84,425	61,888	57,256	70,765	72,012	75,131

Notes: This table shows DID estimates of the effect of universal preschool in different subsamples. The specification includes locality and cohort fixed effects, controlling for parental education, mother's employment and father's earnings (in deciles) at age 2, number of siblings and religion. Mean outcomes of the pre-reform cohorts (born in 1991-1994) in the treatment localities appear in italics. Standard errors (in parentheses) are clustered at the locality level.

Table A9: Robustness Checks - Wild Cluster Bootstrap

Dependent Variable	Main results (1)	Wild cluster bootstrap (2)
High school performance z-score	0.08	0.08
	p-value=0.000	p-value=0.004
	-0.058	*-0.058*
Took the psychometric exam	0.028	0.028
	p-value=0.001	p-value=0.004
	0.389	*0.389*
Postsecondary enrollment by age 19	0.034	0.034
	p-value=0.000	p-value=0.000
	0.157	*0.157*
Any juvenile criminal offense (men)	-0.03	-0.03
	p-value=0.007	p-value=0.008
	0.165	*0.165*
Married by age 21 (women)	-0.017	-0.017
	p-value=0.081	p-value=0.087
	0.318	*0.318*
Number of localities	37	37
Number of observations	84,425	84,425

Notes: Column (1) reports estimated effects and p-values of our main results with clustered standard errors. Column (2) reports p-values from a wild cluster bootstrap estimation to adjust for a small number of clusters. Mean outcomes of the pre-reform cohorts (born in 1991-1994) in the treatment localities appear in italics.

Table A10: Differential Changes in Class Size

	Elementary school	Middle school + high school	Middle school	High school
	(1)	(2)	(3)	(4)
Class size	0.201	-0.100	-0.075	0.462
	(0.402)	(0.384)	(0.596)	(0.426)
	29.361	*30.066*	*33.436*	*27.832*
Number of localities	37	35	32	34

Notes: This table shows DID estimates using average class size as an outcome. The estimation is based on aggregated data at the locality-cohort level. The specification includes cohort and year fixed effects. Mean outcomes of the pre-reform cohorts (born in 1991-1994) in the treatment localities appear in italics. Standard errors (in parentheses) are clustered at the locality level.

Table A11: DID Estimates on Revenues and Expenditures per Capita in Local Authorities

	Revenue per capita (1)	Expenditure per capita (2)	Education expenditure per capita (ages 0-17) (3)
Post X Treatment	-92.386	128.534	68.401
	(149.238)	(164.771)	(157.656)
Outcome Mean	*4160.906*	*4674.353*	*3127.494*
Number of localities	35	35	35
Number of observations	210	210	210

Notes: This table reports DID estimates for revenues and expenditures per capita and education expenditures per capita (ages 0-17). All regressions include locality and year fixed effects. Standard errors, clustered at the locality level, are reported in parentheses. The data are compiled from official reports of the Israeli Central Bureau of Statistics on local authorities for the years 1999-2004. The variables are reported in nominal terms in NIS. One treated and one comparison locality lack financial data in the official reports. Post is a dummy variable that takes the value of one for years 2000-2004.

Table A12: Impact of Universal Preschool on High School Achievement in the Baseline Sample and in an Extended Sample that Includes Late-Treated Localities

Dependent Variable	Baseline Sample (1)	Extended Sample (2)
High school performance z-score	0.08	0.076
	(0.020)	(0.019)
	-0.058	*-0.058*
Graduated from high school	0.028	0.028
	(0.012)	(0.012)
	0.802	*0.802*
Took matriculation exams	0.037	0.043
	(0.011)	(0.010)
	0.763	*0.763*
Matriculation certificate	0.043	0.037
	(0.023)	(0.021)
	0.396	*0.396*
University-eligible certificate	0.035	0.033
	(0.013)	(0.012)
	0.287	*0.287*
4+ English units	0.04	0.036
	(0.016)	(0.014)
	0.364	*0.364*
4+ math units	0.015	0.013
	(0.009)	(0.008)
	0.197	*0.197*
Number of science subjects	0.092	0.082
	(0.041)	(0.037)
	0.688	*0.688*
Number of localities	37	42
Number of observations	84,425	91,193

Notes: This table shows DID estimates of the impact of universal preschool on various educational outcomes. Column (1) displays the estimates for our baseline sample, while Column (2) includes an extended sample of 5 additional localities treated after 2000 (2001-2003). The specification includes locality and cohort fixed effects, controlling for parental education, mother's employment and father's earnings (in deciles) at age 2, number of siblings, and religion. Mean outcomes of the pre-reform cohorts (born between 1991-1994) in the treatment localities appear in italics. Standard errors (in parentheses) are clustered at the locality level. The high school performance z-score (first row), is an average of all standardized individual outcomes.

Table A13: Descriptive Statistics for Pre-reform Cohorts (Treatment and Comparison Localities)

	Main Sample	Siblings Sample	Siblings 'switcher' Sample		Main Sample	Siblings Sample	Siblings 'switcher' Sample
	(1)	(2)	(3)		(1)	(2)	(3)
Panel A: pre-treatment covariates				**Panel B: outcomes**			
Father's years of education	10.54	10.50	10.44	Completed high school	0.84	0.84	0.83
	(3.11)	(3.05)	(3.05)		(0.37)	(0.37)	(0.38)
Mother's years of Education	10.28	10.24	10.04	Participated in the matriculation exams	0.80	0.80	0.79
	(3.00)	(2.89)	(2.88)		(0.40)	(0.40)	(0.41)
Father employed in 1998	0.68	0.69	0.68	Matriculation certificate	0.50	0.49	0.48
	(0.46)	(0.46)	(0.47)		(0.50)	(0.50)	(0.50)
Mother employed in 1998	0.18	0.16	0.15	University-eligible matriculation	0.38	0.37	0.35
	(0.38)	(0.37)	(0.35)		(0.48)	(0.48)	(0.48)
Father's monthly wages in 1998	5,170	5,085	5,203	4+ English units	0.46	0.44	0.43
	(4,003)	(3,659)	(3,846)		(0.50)	(0.50)	(0.50)
Mother's monthly wages in 1998	2,799	2,757	2,817	4+ Math units	0.22	0.21	0.20
	(2,066)	(2,009)	(2,084)		(0.41)	(0.41)	(0.40)
Number of siblings	3.07	3.14	3.32	Number of science subjects	0.55	0.53	0.52
	(1.87)	(1.84)	(1.86)		(0.72)	(0.72)	(0.72)
Share of females	0.49	0.49	0.50	Any juvenile criminal record (men)	0.11	0.11	0.12
	(0.50)	(0.50)	(0.50)		(0.31)	(0.31)	(0.32)
Share of Druze	0.14	0.14	0.15	Participated in the psychometric exam	0.39	0.39	0.38
	(0.35)	(0.34)	(0.36)		(0.49)	(0.49)	(0.49)
Share of Bedouin	0.11	0.11	0.11	Average psychometric score	487	485	484
	(0.31)	(0.31)	(0.31)		(112)	(112)	(112)
				Any postsecondary enrollment	0.29	0.29	0.30
					(0.45)	(0.45)	(0.46)
Number of Localities	37	37	37	Married by age 21 (women)	0.24	0.24	0.25
Number of observations	84,457	69,591	45,684		(0.42)	(0.43)	(0.44)

Notes: This table presents descriptive statistics for individuals' background characteristics and outcomes. The sibling sample includes all individuals that have siblings in the main sample. The siblings "switcher" sample refers to siblings households that have children both in the pre-period cohorts (born in 1991-1994) and a post-reform cohort (born in 1995-1999). Columns 1-3 display the means (and standard deviation (in parentheses)) in each category. Both treatment and comparison localities are included in the sample.

Table A14: Preschool Attendance in Treatment and Never Treated Localities

	Preschool enrollment at age 3	Preschool enrollment at age 4
	(1)	(2)
Father's educ. 12+	-0.018	-0.013
	(0.009)	(0.009)
Mother's educ. 12+	0.012	0.027
	(0.020)	(0.019)
Siblings above median	-0.016	-0.028
	(0.011)	(0.016)
Female	0.001	-0.002
	(0.004)	(0.005)
Treatment x		
Father's educ. 12+	0.021	0.009
	(0.011)	(0.011)
Mother's educ. 12+	0.029	-0.013
	(0.022)	(0.020)
Siblings above median	0.017	0.039
	(0.012)	(0.017)
Female	-0.007	-0.003
	(0.005)	(0.006)
Outcome mean	0.655	0.814
Cohort fixed effect x treatment	Yes	Yes
Locality fixed effect	Yes	Yes
Number of observations	26,204	26,204

Notes: This table reports estimates from a regression where the dependent variable is an indicator for preschool attendance at age 3 (column (1)) and age 4 (column (2)) and the explanatory variables are family background characteristics and child gender. The models also include interactions between these covariates and a treatment indicator, locality fixed effects, and cohort fixed effects interacted with a treatment indicator. The sample includes treated and never treated localities. Enrollment data is from the post-reform period.

Table A15: Heterogeneous Effects of Universal Preschool by Predicted Likelihood of Matriculation

	High school z-score	Postsecondary enrollment by age 19	Took the psychometric exam	Any juvenile criminal record (men)	Married by age 21 (women)
	(1)	(2)	(3)	(4)	(5)
Exposure to Preschool	0.034	0.049	0.015	-0.010	-0.006
	(0.019)	(0.014)	(0.014)	(0.015)	(0.012)
Exposure to Preschool X	0.057	-0.027	0.015	-0.016	-0.011
Low/Median Predicted Outcome	(0.021)	(0.015)	(0.015)	(0.015)	(0.019)
Mean outcome	-0.058	0.157	0.389	0.165	0.318
Number of observations	84,425	74,424	84,457	43,345	31,256
Number of Localities	37	37	37	37	37

Notes: This table shows the estimated effect of universal preschool allowing for heterogeneity of the effect by including the main treatment indicator (Exposure to Preschool) and its interaction with the dummy variable, Low/Median Predicted Outcome. The regression is fully saturated: all the control variables are also interacted with this dummy variable. The specification includes locality and cohort fixed effects, controlling for parental education, mother's employment and father's earnings (in deciles) at age 2, number of siblings and religion. Mean outcomes of the pre-reform cohorts (born in 1991-1994) in the treatment localities appear in italics. Standard errors (in parentheses) are clustered at the locality level.

Table A16: Heterogeneous Effects of Universal Preschool by Predicted Likelihood of Matriculation

	Predicted Likelihood of Matriculation		
	Low	Medium	High
Dependent Variable	(1)	(2)	(3)
High school performance z-score	0.084	0.103	0.034
	(0.029)	(0.028)	(0.023)
	-0.447	*0.030*	*0.586*
Graduated from high school	0.035	0.025	0.006
	(0.024)	(0.012)	(0.006)
	0.648	*0.888*	*0.974*
Took matriculation exams	0.058	0.031	0.006
	(0.020)	(0.012)	(0.006)
	0.584	*0.861*	*0.965*
Matriculation certificate	0.038	0.075	0.017
	(0.025)	(0.034)	(0.021)
	0.202	*0.436*	*0.728*
University-eligible certificate	0.034	0.06	0.018
	(0.014)	(0.017)	(0.017)
	0.114	*0.294*	*0.629*
4+ English units	0.038	0.07	0.008
	(0.016)	(0.020)	(0.021)
	0.160	*0.384*	*0.749*
4+ math units	0.014	0.019	0.004
	(0.007)	(0.009)	(0.022)
	0.077	*0.181*	*0.472*
Number of science subjects	0.058	0.113	0.085
	(0.035)	(0.054)	(0.051)
	0.358	*0.734*	*1.280*
Took the psychometric exam	0.019	0.041	0.012
	(0.010)	(0.012)	(0.016)
	0.183	*0.430*	*0.742*
Postsecondary enrollment by age 19	0.016	0.033	0.045
	(0.006)	(0.010)	(0.012)
	0.068	*0.149*	*0.352*
Any juvenile criminal offense (men)	-0.019	-0.034	-0.027
	(0.013)	(0.012)	(0.010)
	0.195	*0.164*	*0.097*
Married by age 21 (women)	-0.006	-0.018	-0.022
	(0.016)	(0.016)	(0.021)
	0.393	*0.292*	*0.151*

Notes: This table shows the estimated effect of universal preschool, by tertiles of predicted matriculation eligibility defined by the pre-reform relationship between matriculation eligibility and background characteristics. The specification includes locality and cohort fixed effects, controlling for parental education, mother's employment and father's earnings (in deciles) at age 2, number of siblings and religion. Mean outcomes of the pre-reform cohorts (born in 1991-1994) in the treatment localities appear in italics. Standard errors (in parentheses) are clustered at the locality level.

Table A17: Treatment effects on quality of high-school enrolled

Dependent Variable	
High school quality (matriculation eligibility rate of tenth graders in 2008)	0.010
	(0.007)
	0.453
	N=71,453
Probability to have missing data on high school quality (no HS, or HS established after 2009)	0.014
	(0.039)
	0.134
	N=84,425
Number of localities	37

Notes: This table shows estimates of the effect of universal preschool on the quality of the high school attended by the student, proxied by the share of tenth graders eligible for a matriculation certificate in the pretreatment year (2008). The specification includes locality and cohort fixed effects, controlling for parental education, mother's employment and father's earnings (in deciles) at age 2, number of siblings and religion. Mean outcomes appear in italics. Standard errors (in parentheses) are clustered at the locality level.

Table A18: Effects of Universal Preschool on Maternal Employment
Mothers of the individuals included in our study (1991-1999 cohorts)

Dependent Variable	All Mothers (1)	Mother's Years of Education<12 (2)	Mother's Years of Education>=12 (3)
Mother employed (age 3)	0.010	0.011	0.015
	(0.008)	(0.009)	(0.013)
	0.163	*0.099*	*0.334*
Mother employed (age 4)	0.007	0.006	0.018
	(0.009)	(0.009)	(0.013)
	0.169	*0.105*	*0.342*
Mother employed (age 5)	0.013	0.012	0.020
	(0.009)	(0.009)	(0.014)
	0.174	*0.106*	*0.358*
Mother's months worked (age 3)	0.028	0.041	0.147
	(0.079)	(0.066)	(0.136)
	1.294	*0.633*	*3.084*
Mother's months worked (age 4)	0.024	0.018	0.159
	(0.086)	(0.065)	(0.148)
	1.367	*0.682*	*3.222*
Mother's months worked (age 5)	0.048	0.008	0.234
	(0.084)	(0.071)	(0.150)
	1.430	*0.726*	*3.337*
Mother's log annual wages (age 3)	0.033	0.020	0.041
	(0.049)	(0.083)	(0.064)
	8.932	*8.238*	*9.486*
Mother's log annual wages (age 4)	0.033	0.041	0.021
	(0.039)	(0.066)	(0.048)
	9.173	*8.491*	*9.732*
Mother's log annual wages (age 5)	-0.017	-0.071	0.016
	(0.048)	(0.069)	(0.057)
	9.375	*8.746*	*9.877*
Number of localities	37	37	37
Number of observations	84,367	50,724	33,643

Notes: This table shows DID estimates of the impact of exposure to the Preschool Law on maternal employment when the child was 3-5 years old. The basic unit of observation is children in the years before and after implementation of universal preschool (born in 1991-1999). The specification includes locality and cohort fixed effects, controlling for parental education, number of siblings, and religion. Mean outcomes of the pre-reform cohorts (born in 1991-1994) in the treated localities appear in italics. Standard errors (in parentheses) are clustered at the locality level.

Table A19: Effects of Universal Preschool on Maternal Employment

Panel data of mothers living in the localities of the study, 1995-2004

	All		Years of Education<12		Years of Education>=12	
Dependent Variable	Mothers of Children Aged 3-5 (1)	Other Mothers (2)	Mothers of Children aged 3-5 (3)	Other Mothers (4)	Mothers of Children aged 3-5 (5)	Other Mothers (6)
Employed	0.003	0.006	0.010	0.008	0.004	0.013
	(0.007)	-0.007	(0.007)	(0.008)	(0.011)	(0.011)
	0.175	*0.202*	*0.106*	*0.121*	*0.354*	*0.379*
Number of months worked	-0.070	-0.018	-0.012	-0.015	0.005	0.079
	(0.068)	(0.056)	(0.057)	(0.070)	(0.104)	(0.092)
	1.435	*1.631*	*0.702*	*0.828*	*3.349*	*3.384*
(Log) Annual real wages	-0.034	0.015	-0.053	-0.006	-0.032	0.012
	(0.038)	(0.027)	(0.055)	(0.041)	(0.040)	(0.026)
	9.214	*9.234*	*8.508*	*8.654*	*9.758*	*9.638*
Number of localities	37	37	37	37	37	37
Number of observations	216,596	206,275	125,930	114,719	90,666	91,556
Number of observations with positive wages	54,874	60,109	17,589	18,051	37,285	42,058

Notes: This table shows DID estimates of the impact of the Preschool Law on mothers who live in the localities of the of the study's main sample . The basic unit of observation is the mother-year level. The specification includes locality and year fixed effects, controlling for education, age, age squared and religion. Mean outcomes in the pre-treatment years (1995-1999) in the treated localities appear in italics. Standard errors (in parentheses) are clustered at the locality level.

Table A20: Effect of the Preschool Law on Preschool Enrollment at the Locality Level

	Age 3	Age 4	Age 5
	(1)	(2)	(3)
	A. All Arab Localities		
Preschool Law exposure	0.603	0.555	0.009
	(0.050)	(0.051)	(0.033)
Number of localities	52	52	52
	B. Localities of the Study		
Preschool Law exposure	0.597	0.492	-0.043
	(0.056)	(0.062)	(0.026)
Number of localities	36	36	36

Notes: This table shows DID estimates of the impact of the Preschool Law on preschool enrollment at different ages. The estimation is based on aggregated data at the locality-year level weighted by population size. The specification includes locality and year fixed effects. Standard errors (in parentheses) are clustered at the locality level.

Table A21: Sources of Estimates Reported in Table 9

Study	References for estimates	References for counterfactual mode of care	References for maternal education
Gray-Lobe et al. (2023)	Table IV, column (2); Table III, column (8)	Section III.B, p. 379	n.a.
Havnes and Mogstad (2011)	Table IV, columns (1) and (4)	Section VIII, p. 124	Table 3
Deming (2009)	Table V, column (1)	n.a.	Table 1, family fixed effects subsample. Presents means separately for black and white samples. Our weighted calculation for the entire population.
Bailey et al. (2021)	Table I, columns (1) and (6)	Section IV, p. 3978, and Table 1, p. 3981	Section IV, footnote 22, p. 3977
This study	Estimates for high school graduation are based on Table 3, column (1). Estimates for college enrollment are based on Table 5 column (1)under “Enrolled at academic institution.” Both are inflated by the increase in preschool enrollment (0.6).	See discussion in background section	Table A3
Belfield et al. (2006)	Table 1 and authors' calculations.	Section III-a, p. 1481, in Garcia et al. (2023), which studies the same project.	Heckman et al. (2010), Table 1
Campbell et al. (2012)	Table 3 and discussion on p. 10. Findings refer to earning a Bachelor's degree rather than college enrollment.	Section 2.2, p. 13 in Garcia et al. (2018), which studies the same project.	Table 2
Heckman et al. (2010)	Table III, columns (2) and (3); Table V, columns (2) and (3)	Section III-a, p. 1481 in Garcia et al. (2023), which studies the same project.	Table 1
Anderson (2008)	Table 6, columns (3)-(4) and (8)-(9)	Section 2.2, p. 13 in Garcia et al. (2018), which studies the same project.	Campbell et al. (2012), Table 2
Elango et al. (2016)	Figure 4.6.		

Notes: This table presents the sources of the estimates reported in Table 9.

Table A22: Pre-reform and Postreform Cohorts of the Study, by Age at Observation in Their Locality of Residence

Year of birth	Type of cohort	Source	Age observed in their locality of residence
1991	pre-reform	Israeli Census of Population 1995	4
1992	pre-reform	Israeli Census of Population 1995	3
1993	pre-reform	Israeli Census of Population 1995	2
1994	pre-reform	Israeli Census of Population 1995	1
1995	post-reform	Israeli Census of Population 1995/ Israeli Registry of Citizens 2000	0 / 5
1996	post-reform	Israeli Registry of Citizens 2000	4
1997	post-reform	Israeli Registry of Citizens 2000	3
1998	post-reform	Israeli Registry of Citizens 2000	2
1999	post-reform	Israeli Registry of Citizens 2000	1

Notes: This table details the age each cohort was observed in their locality of residence. For the 1995-born cohort, the selected locality of residence was the one observed in the 1995 Census, unless the individual was not yet born at the time of the census. In the latter case, we report the locality recorded in 2000.