I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 18037

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

Elad DeMalach
Analia Schlosser

JULY 2025

IZA Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 18037

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

Elad DeMalach
Bank of Israel and Tel Aviv University

Analia Schlosser
Tel Aviv University, CEPR, CESifo, and IZA

JULY 2025

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

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

We estimate the causal impacts of universal preschool by leveraging a quasi-experimental design based on Israel's implementation of free public preschool for children ages 3 and 4 beginning in September 1999. We focus on the Arab population, who were the main beneficiaries of the first phase of the Law's implementation. Using a difference-in-differences research design, we find that universal preschool enhanced individuals' academic performance from elementary school through high school, improved the learning environment, and increased postsecondary enrollment. Additional benefits include reduced juvenile delinquency among males and decreased early marriage among females.

JEL Classification: I24, I25, J20

Keywords: preschool, early childhood, education, minorities

Corresponding author:
Analía Schlosser
Tel Aviv University
P.O.B. 39040
Ramat Aviv, Tel Aviv, 69978
Israel
E-mail: analias@tauex.tau.ac.il

* We thank seminar participants at Reichman University, Bar Ilan University, Tel Aviv University, Hebrew University of Jerusalem, the Geneva School of Economics, Bank of Israel and Management and participants at the Annual Conference of the Israeli Economic Association, the Early Childhood Education Conference at the Taub Center for Policy Research, SOLE meetings, the NBER SI Children's Meeting, the IZA/ECONtribute in Economics of Education, the CESIfo Venice Summer Institute, and the CESIfo Area Conference in Economics of Education. We thank Tatiana Baron for her contribution at the early stage of the project. We thank Avigail Sageev for her research assistance with the PISA data, Rabab Hijazi for her fantastic work documenting stories from the field about the law implementation, and Fatma Kassem and Nabila Espanioly for sharing with us their knowledge on early childhood education within Arab society. Research was conducted in the research room of the Central Bureau of Statistics based on de-identified individual record files prepared by the Central Bureau of Statistics. This research was supported by Israeli Science Foundation grant no. 1929/19. Schlosser gratefully acknowledges the financial support of the Center for AI and Data Science, the Foerder Institute for Economic Research, and the Pinhas Sapir Center for Development at Tel Aviv University.

1. Introduction

Educational interventions at young ages can have substantial long-term effects on adult outcomes (Heckman and Masterov, 2007; Cunha and Heckman, 2007; Currie and Almond, 2011; Heckman et al., 2013). These findings have spurred growing policy interest in public preschool as a tool to reduce income inequality and promote intergenerational mobility.[1] Most European countries—including the U.K., France, Germany, all Nordic nations—and several U.S. states and cities now offer universal or large-scale preschool programs aimed at enhancing children's social and cognitive development. Yet, evidence on the causal impacts of such universal programs remains scarce due to identification challenges and the limited availability of long-term follow-up data.

This paper examines the causal effects of universal preschool by leveraging a quasi-experimental design based on Israel's implementation of free public preschool for children ages 3 and 4 beginning in September 1999. We offer a unique causal analysis of the life-cycle impacts of universal preschool, drawing on multiple datasets that track individuals for up to 20 years after treatment. We follow students through elementary, middle, and high school, analyzing test scores, matriculation exam success, and postsecondary enrollment. We also evaluate the impact on two key social outcomes: juvenile crime and early marriage. Finally, we investigate potential mechanisms, including elementary and middle school learning environments and changes in maternal employment and income.

We focus on a particularly disadvantaged group in Israeli society: the Arab population living in cities, towns, and villages (hereafter localities) with low socioeconomic status. Research consistently shows that disadvantaged children benefit more from public preschool, largely due to lower-quality home environments and alternative childcare (van Huizen and Plantenga, 2018).

Our identification strategy relies on variation created by the initial rollout of Israel's Compulsory and Free Preschool Law for Ages 3 and 4, enacted in September 1999. Although the law required free preschool for all children at these ages, implementation began in only the most disadvantaged localities—predominantly Arab—resulting in a sharp expansion in public preschool access and enrollment in these treatment localities.

We exploit this targeted rollout using a difference-in-differences (DID) design focusing on Arab localities. Specifically, we compare changes in outcomes across treated and untreated cohorts in treated localities to similar cohorts in Arab localities that remained unaffected during the first phase of the law

[1] See, e.g., President Obama's 2013 State of the Union Address: https://obamawhitehouse.archives.gov/the-press office/2013/02/12/remarks-president-state-union-address, and President Biden's The American Families Plan: https://www.whitehouse.gov/briefing-room/statements-releases/2021/04/28/fact-sheet-the-american-families-plan/.

implementation. To support our identification strategy, we conduct several robustness tests to rule out confounding trends, shifts in sample composition, or school inputs. We also validate our results using a family fixed effects approach, comparing exposed and unexposed siblings within treatment and comparison localities.

We find that universal preschool had a profound impact on the public preschool enrollment of Arab children in treatment localities. Enrollment at age 4 rose from 23% to 90%, and at age 3 from 16% to 80%, while remaining relatively stable in comparison localities. The reform also significantly improved educational attainment: high school graduation, matriculation exam participation and pass rates, and the quality of matriculation certificates (reflected in more credit units in math, English, and science) all increased. We further observe higher participation and scores in the psychometric college-entrance exam (Israel's SAT-equivalent), along with increased postsecondary enrollment in both academic and vocational institutions.

These long-term gains appear partly driven by early improvements in native language and math proficiency, and by a more supportive school climate, as students report better relationships with teachers and peers and a stronger sense of security in school. Notably, we find no corresponding rise in maternal employment or income, indicating that the educational effects are not explained by changes in family economic conditions. Importantly, the benefits extend beyond education. Boys exposed to universal preschool were less likely to have a juvenile criminal record, and young women were more likely to delay marriage. These findings are especially meaningful given the high crime rates among Arab youth and early marriage norms for women in this traditional community.

The literature on early childhood education has grown substantially in recent years (see van Huizen and Plantenga, 2018; Cascio, 2023; Duncan et al., 2023; Bruhn and Emick, 2023, for recent reviews). Early work focused on small-scale programs from the 1960s and 1970s targeting disadvantaged children with intensive services through randomized control trials (e.g., Schweinhart et al., 2005; Anderson, 2008; Heckman et al., 2010, 2013). These studies found significant improvements in cognitive and non-cognitive skills across the life course. The prevailing view highlights short-term gains that sometimes fade in school years but reappear later in life. However, due to the selective targeted populations, intensive services, and historical context, these results may not generalize to universal preschool. Moreover, such targeted programs face challenges in scalability due to high costs and the difficulty of maintaining individualized treatment and quality at scale.

Evidence on universal or large-scale preschool remains relatively limited, primarily due to difficulties in identifying causal effects. Most studies focus on specific periods, such as early childhood outcomes

(Berlinski et al., 2008, 2009; Fitzpatrick, 2010; Cornelissen et al., 2018; Cascio, 2023) or long-term results like high school graduation, college attendance, and employment (Havnes and Mogstad, 2011, 2015; Bailey et al., 2021). Only a few examine both short- and long-term outcomes (Felfe et al., 2015; Blanden et al., 2016; Gray-Lobe et al., 2023), and findings remain mixed. These inconsistencies likely stem from variation in counterfactual care, family backgrounds, preschool quality, and whether early gains are sustained through elementary education—factors not always observable in the data.

Our paper contributes to the early childhood education literature in several ways. First, we offer a detailed causal analysis of the life-cycle effects of large-scale universal preschool by integrating multiple data sources and tracking individuals for up to 20 years without attrition. We examine a broad set of outcomes—educational performance, postsecondary attainment, juvenile crime, and early marriage—which enables us to assess heterogeneous effects across individuals and outcome margins.

Second, our setting provides a clear counterfactual: absent the reform, most children would have stayed at home receiving parental care. This differs from other contexts where universal preschool often replaces private care for high-SES families or existing early childhood programs for disadvantaged groups (e.g., Kline and Walters, 2016). In addition, the reform led to near-universal preschool attendance from a low baseline, allowing us to study the effects of aligning all children at the starting line of formal education. Third, we examine potential mechanisms by documenting short- and medium-term effects on learning environment and maternal labor supply and income.

Our study is also the first large-scale analysis of the long-term causal impacts of preschool among Arabic speakers—the world's fifth most spoken language. This is important due to the diglossic nature of Arabic: the formal language used in school (Modern Standard Arabic - MSA) differs from the spoken language, creating challenges for literacy acquisition (Ferguson, 1959; Saiegh-Haddad, 2003). Notably, a World Bank report (Gregory et al., 2021) highlights severe learning poverty in the MENA region, where 59% of children cannot read an age-appropriate text by age 10. While preschool is key to building early literacy, the region has among the lowest pre-primary enrollment globally. More broadly, our findings are also relevant for populations facing linguistic gaps between home and school, where early exposure to formal grammar structures and academic vocabulary can significantly enhance school readiness and later academic achievement.

The remainder of the paper is organized as follows. Section 2 provides background on early education in the Israeli Arab population and the implementation of the law. Section 3 outlines our identification strategy, and Section 4 describes the data and presents summary statistics. Section 5 reports the main results. Section 6 discusses falsification and robustness tests and presents family fixed effect estimates.

Section 7 explores heterogeneity across subgroups and examines potential mechanisms underlying the long-term effects by analyzing intermediate outcomes. Section 8 compares our findings to other early childhood programs worldwide and to later-stage educational interventions in Israel. Section 9 concludes.

2. Institutional Background

The Arab population comprises 21% of Israel's citizens, reaching 2 million by the end of 2021. Compared to the Jewish population, they have lower educational attainment, lower incomes, and higher poverty rates (Bank of Israel, 2021). About 84% are Muslim, with Christian (7%) and Druze (8%) communities making up the remainder.[2] Arab society is generally traditional, especially regarding gender roles. The population is largely residentially segregated: 85% live in Arab localities (where they form nearly the entire population), 10% in mixed Arab-Jewish localities, and 5% are Bedouins living in places that have not been officially recognized by the Ministry of Interior.[3]

The Arab and Jewish education systems are separated through high school. Most Arab students attend Arab public schools, taught mainly by Arab staff. The language of instruction in Arab schools is Modern Standard Arabic (MSA), which differs from the colloquial spoken Arabic, creating challenges for literacy acquisition (Saiegh-Haddad, 2003; Abu Ahmad et al., 2014). Research highlights the importance of early exposure to MSA in preschool for building foundational literacy skills (Abu-Rabia, 2000; Aram et al., 2013; Saiegh-Haddad and Spolsky, 2014; Saiegh-Haddad, 2022).[4]

Unlike Jewish children, who had high preschool enrollment in the 1990s, only a small share of Arab children attended preschool prior to implementation of the 1999 Preschool Law. In the 1998/1999 school year, enrollment for Arab 3- and 4-year-olds was 21.3% and 32.2%, compared to 79.7% and 89.1% among Jewish children. Private preschool enrollment was also much lower among Arab children (1.2% and 1.4% for ages 3 and 4) than among Jewish children (9.6% and 3.8%) (CBS, 2000).

Enrollment among 5-year-old Arab children was relatively high at 81%, though still 12 percentage points below that of Jewish children (CBS, 2000). This higher rate is largely due to the 1949 Compulsory Schooling Law, which mandated free public preschool for age 5. In contrast, until 2000, the provision of preschool for ages 3 and 4 was left to local authorities, who were not legally required to offer such

---

[2] The data is from 2020. The authors' calculations are based on Table 2.3 in CBS (2021b).
[3] The authors' calculations are based on Table 1.2 in the Inaugural Annual Statistical Report on Arab Society in Israel, published by the Israel Democracy Institute (2022). East Jerusalem is not included.
[4] See Saiegh-Haddad et al. (2022) for research on diglossia and dialectal variation in literacy development. In the U.S. context, several studies suggest that the linguistic distance between African American English and Mainstream American English (also termed "Standard American English") may contribute to the achievement gap and school readiness between African American and White students (see, e.g., Brown et al., 2015; Terry et al., 2022).

services. Although the Ministry of Education subsidized 80%–90% of preschool costs for new immigrants and residents in designated development areas, most Arab localities did not qualify.[5] Combined with the limited financial resources of Arab local governments, this resulted in little to no provision of preschool services (Abu-Jaber, 1994; Israel State Comptroller, 1992). For instance, in 1993, only 15 out of 100 Arab local authorities surveyed by Ghanem (1993) offered such services.

As a result, most Arab children under age 5 stayed at home and did not attend any form of daycare. According to the 2009 PISA Student Questionnaire, only 34% of Arab children from the 1993 cohort attended any formal pre-primary program for more than a year, compared to 86% of Jewish children. At the time, Arab women's labor force participation was also extremely low—17% in 1998, compared to 64% among Jewish women.[6]

In September 1999, the Israeli government began implementing the Compulsory and Free Preschool Law for Ages 3 and 4, mandating that the State provide free preschool education to all Israeli children in this age group.[7] While the State was obligated to provide universal preschool, parents were not required to enroll their children (see Kimhi, 2012; protocol of the Knesset's Education Committee, May 29, 1984). Implementation began in the most disadvantaged localities and was intended to expand annually to cover the entire country within ten years, based on socioeconomic clusters ranked from 1 (lowest) to 10 (highest).[8]

Starting in September 1999, localities classified in clusters 1 and 2 received universal preschool access. As 91% of these localities were Arab, the law significantly impacted the Arab population, granting preschool access to most Arab children for the first time. The law also applied to designated development areas, though these localities had already received substantial preschool subsidies since the mid-1980s, making preschool effectively universal there prior to the reform. Preschools operated five days a week, 6.5 hours per day, with class sizes of up to 35 children. Each classroom was staffed by one certified teacher and one teacher aide. Teachers were required to hold certification from an academic institution

[5] These include localities with the status of "National Priority," "Confrontation Line," and neighborhoods and localities included in the "Urban Renewal Project." Preschool subsidies in development-targeted areas began in 1978 (*Ma'ariv*, June 4, 1978), but Arab localities were excluded until the mid-1980s. Some were later included—for example, Government Decision 323 (April 1987) extended eligibility to Druze localities, and another decision equalized benefits for Arab and Jewish localities near Israel's border (11th Knesset Proceedings, July 6, 1988, p. 3591; 12th Knesset Proceedings, Jan. 21, 1991, p. 2064).

[6] The authors' calculations are based on data from the 1998 CBS Labor Force Survey.

[7] For reviews of the law's implementation, see Kop (2002) and Blass and Adler (2004).

[8] The Central Bureau of Statistics assigns each locality a socioeconomic index based on factors like income, housing, education, and employment. Localities are ranked and grouped into 10 clusters that are internally similar based on this index. See CBS (2003) for details.

recognized for training teaching staff in Israel, while teacher aides needed at least 12 years of schooling and a teaching aide certificate. Some preschools also employed student teachers from early childhood education colleges as part of their practical training requirements. During the study period, the average class size in our sample was 32. Appendix A provides additional details on the pedagogical approach and implementation.

The government initially intended to expand the law's coverage gradually based on each locality's cluster classification. In practice, however, this rollout was repeatedly delayed due to budget constraints, and only in 2015—fifteen years later—was nationwide coverage officially achieved.[9]

Figure 1 shows the geographical distribution of Arab localities. Those that gained preschool access under the first phase of the law are marked in red; all others are marked in blue. In the Central district, no Arab localities were included in the first phase. The Southern district consists entirely of Bedouin localities that were included but differ in many respects from other Arab communities (see Abu-Bader and Gottlieb, 2013). The Northern district is unique in containing both treated and untreated Arab localities, making it the focus of our analysis.

3. Research Design and Identification Strategy

Our analysis focuses on Arab localities in Israel's Northern district, including 15 localities that first received universal free preschool in September 1999 (treatment localities) and 22 that did not experience major changes in access during the initial phase of the law implementation (comparison localities). In the comparison group, 17 localities had received subsidies prior to the law ("always treated"), while 5 had no public preschool access during the period of study: 1999-2003("never treated").[10]

Figure 2 shows enrollment rates in public preschools by age and year, distinguishing between treated, never-treated, and always-treated localities included in our analysis sample. To simplify the discussion and following Ministry of Education convention, we define the law's first year as 2000 (1999/2000 school year). Enrollment in treated localities rose sharply—from 18% to 91% at age 3, and from 31% to 93% at age 4—between 1999 and 2003. In contrast, enrollment rates in comparison localities remained largely

[9] Some localities became eligible a few years after 2000 due to reclassification into clusters 1 or 2. We exclude them from the main analysis, as post-high school outcomes and post-reform cohorts are not fully observable. Appendix D shows that including them yields similar results.

[10] We include only localities with independent local authorities that have their own CBS-defined socioeconomic clusters. We exclude 5 localities added to the law later due to reclassification, 3 Druze localities in the Golan Heights without 1995 census data (and thus no CBS cluster), and 6 localities that could not be classified as treatment or comparison groups.

unchanged. Age 5 enrollment was already close to 100% throughout and showed no clear trend. Patterns in the Northern district closely mirror those in all Arab localities (see Figure A1).

To assess the impact of universal preschool education, we use a difference-in-differences (DID) approach. Specifically, we compare changes in outcomes between cohorts of children in treatment and comparison localities who reached preschool age before and after the Preschool Law. Pre-reform cohorts were born in 1991–1994, and post-reform cohorts in 1995–1999, as the law took effect in the 1999/2000 school year.

The estimating equation for our analysis is:

$$Y_{ist} = \alpha + \beta Exposed\_Preschool_{s(t+4)} + \gamma X_{ist} + \delta_s + \lambda_t + \varepsilon_{ist} \tag{1}$$

where Y_{ist} is the outcome for individual *i*, residing in locality *s,* born in year t. $Exposed\_Preschool_{s(t+4)}$ equals 1 if the child lived in a treatment locality and was no older than 4 when the law was implemented, and 0 otherwise. X_{ist} includes the following individual controls: parental education, maternal employment at child age 2, paternal income decile at age 2 (plus a dummy for missing/zero income), number of siblings, religion (Muslim, Christian, Druze), and gender.[11] δ_s are locality fixed effects that control for any cohort-invariant differences across localities, and λ_t are cohort fixed effects that nonparametrically control for time effects at the cohort level. Standard errors are clustered at the locality level. The coefficient of interest β captures the intention-to-treat (ITT) effect of universal preschool. This is the parameter of interest from a policy perspective when the objective is to capture the effect of universal preschool education. In Section 8, we also report local average treatment effect (LATE) estimates, obtained by scaling the ITT by the increase in enrollment due to the reform, to facilitate comparison with prior studies.

Our empirical strategy relies on the assumption that, absent the law, trends in outcomes would have been similar in treatment and comparison localities. It also requires that no shocks or policies during the study period differentially affected children aged 4 or younger relative to those aged 5–9 in treated versus comparison localities. This assumption is plausible, as no other reforms specifically targeted young children in treated or comparison localities during this time. Moreover, local shocks are unlikely to differentially affect such narrow age bands, which typically share similar family environments and

[11] We define employment as earning at least half the minimum wage, though results are robust to a lower threshold of any positive earnings. Given the very low labor force participation of Arab women during the sample period, we control for maternal employment rather than maternal wage deciles.

institutions. To validate our design, we conduct a range of robustness tests summarized in Section 6 and detailed in Appendix D.

While our DID is not staggered, one concern is that always-treated localities may exhibit dynamic effects that could bias our results (see Callaway and Sant'Anna, 2021; Roth et al., 2023). However, as these localities began receiving preschool services in the mid-1980s—over a decade before the reform—we expect treatment effects to be stable over our study period. This is supported by Figure 2, which shows stable enrollment rates in always-treated localities. As an additional check, we conduct robustness tests in Appendix D using only never-treated or always-treated localities as the comparison group, with results consistent with our main estimates.

4. Data and Descriptive Statistics

Data

Our dataset was created by merging administrative records from multiple sources stored in the research rooms of the Israel Central Bureau of Statistics. The starting point is the Israeli population registry, from which we selected all Israeli Arabs born in 1991–1999. The registry also includes information on gender, marital status, and locality of residence during childhood. Using personal identifiers, we merge these data with the Israeli educational registry, which provides information on individuals' enrollment in primary, secondary, and postsecondary education.

We then merge the data with student records on standardized exams administered by the Ministry of Education (MOE). The first set is the GEMS (Growth and Effectiveness Measures for Schools, or *Meitzav* in Hebrew) exams, conducted annually in the fifth and eighth grades in a representative sample of schools in four subjects: native language (i.e., Arabic), English, math, and science. These exams also include a questionnaire on the learning environment filled out by students in these schools in grades 5–9.

We also merge student data from matriculation exams, i.e., national high school exit exams taken in core and elective subjects in grades 10–12. The matriculation certificate is a prerequisite for postsecondary admission to Israeli universities and academic colleges. We also obtain information on student performance on the psychometric exam, a standardized test (similar to the U.S. SAT) used together with the matriculation certificate as the main admission criterion in higher education. Finally, we merge our dataset with administrative police records on juvenile crimes, which indicate whether an individual was arrested and had a criminal record in youth (until age 18) and the general category of the crime.

We further enrich the student data by adding family background characteristics: parental education from the education registry and number of siblings from the population registry. In addition, we use administrative records from the Israel Tax Authority to obtain information on the employment and earnings of the parents of the individuals in the main sample. Given that, at the time of dataset construction, such information was only available up to 2018, we cannot analyze the employment and earnings of the cohorts affected by the reform, as they are still too young.

Our final sample includes 84,425 individuals from the treatment and comparison localities in the relevant cohorts. In Appendix B we provide further information on the data and sources. Table A1 places the study outcomes on an age timeline, providing an overview of the cohorts and time horizon covered. Table A2 provides a full description of the outcome variables and their definitions.

Descriptive Statistics

Table A3 presents the socioeconomic characteristics of the treatment and comparison localities measured in 1999, prior to the law's implementation.[12] In column (3) of the table we report differences between the two groups of localities. The population in treatment localities was significantly more disadvantaged along several dimensions than the population in comparison localities. For example, income per capita was about 16% lower, the dependency ratio was higher, and educational attainment was lower. This is unsurprising as the law was first implemented in the two lowest socioeconomic clusters of localities. Notably, the treatment and comparison localities are similar in terms of average population size.

Panel A of Table A4 presents family background characteristics of the children in the pre-reform cohorts (born in 1991–1994) in the treatment and comparison localities. These, too, show a more disadvantaged treatment population. The parents of children in treatment localities were less educated, had a lower income, and had more children. There is also a different ethnic composition: the Druze Arabs are in the comparison localities, while the Bedouin are mostly in treatment localities. In Panel B of Table A4 we examine differences in outcomes of the individuals in the pre-reform cohorts between treatment and comparison localities. Most point to the relative advantage of the population in the comparison localities during the pre-reform period.

[12] Data for the locality profiles were compiled by the Central Bureau of Statistics, based on administrative records from 1999 (CBS, 2003).

5. Results

High School Outcomes

Table 1 displays our main DID estimates from equation (1) for high school outcomes for the full sample (column (1)) and by gender (columns (2) and (3)). We also report mean outcomes (in *italics*) of the pre-reform cohorts in treatment localities. To deal with multiple hypothesis testing, summarize our high school outcomes, and increase statistical power, we present treatment effects on an index of high school performance (at the top of the table), computed as the standardized average (z-score) of all standardized individual high school outcomes.

We find that implementation of the Preschool Law significantly improved high school graduation and matriculation exam outcomes of Israeli Arabs in treatment localities. Universal preschool increased the likelihood of graduating from high school by 2.8 percentage points (a 3.5% increase relative to the pre-reform mean) and increased participation in the matriculation exams by 3.7 percentage points (5%). The likelihood of obtaining a matriculation certificate rose by 4.3 percentage points (11%) and the probability of obtaining a matriculation certificate that meets university entrance requirements increased by 3.3 percentage points (12%).[13] Improvement in the quality of the matriculation certificate is also reflected in the increased share of individuals who earned at least 4 credit units in English and math (11% and 8%, respectively). Furthermore, the number of science subjects attained in the matriculation certificate increased by 0.9 (13%).[14] While some estimates for the effect of universal preschool differ by gender, the general pattern points to a significant improvement in high school outcomes for both boys and girls.

Figure 3 presents estimates and 95% confidence intervals for the high school performance index and all high school outcomes in the form of an event-study design where year zero denotes the first year of the law's implementation.[15] The figure also reports p-values for a joint test of significance for the coefficients in the pre- and post-reform periods. Estimates of the pre-reform period are small in

---

[13] A matriculation certificate that meets university entrance requirements needs to include at least 4 credit units in English and a passing grade in math at the 3-unit credit level.

[14] Science subjects include physics, chemistry, biology, and computer science.

[15] The figure plots estimates for β_τ and their standard errors from the following model:

$$Y_{ist} = \alpha + \sum_{\tau=-4,\tau\neq-1}^{\tau=4} \beta_\tau Treated_s \times D_{i,2000+\tau} + \gamma X_{ist} + \delta_s + \lambda_t + \varepsilon_{ist}$$

where, for a given τ, the indicator $D_{2000+\tau}$ takes a value of 1 if the individual was 4 years old in year 2000+τ, and 0 otherwise. The omitted period is $\tau = -1$, which is the year before the law's implementation. For $\tau = -4, \dots, 4$, β_τ denotes the evolution of outcomes in treatment localities net of equivalent changes in comparison localities. In Figure A2 we plot the respective unconditional outcome means for the cohorts of the sample living in treated, always-treated, and never-treated localities.

magnitude and not statistically different from zero (individually or jointly), and they do not show any clear pattern of a differential trend in outcomes in treatment versus comparison localities before implementation of the law. This is also consistent with the placebo exercise we discuss in Table A7 (Section 6), where we find no differential changes in outcomes between treatment and comparison localities when we compare the first two and last two years of the pre-reform period. By contrast, the post-reform period estimates observed in Figure 3 show a substantially greater increase in outcomes for the treated localities than the comparison localities for the cohorts exposed to universal preschool education relative to the pre-reform period.

We assess the robustness of this result to possible violations of the parallel trends assumption, using the method suggested by Rambachan and Roth (2023) and focusing on the high school performance z-score. Results, discussed in Appendix C and reported in Figure A3, suggest that our results would remain significant even if we allow for some deviations from the parallel trends assumption.

Postsecondary Outcomes

Psychometric Exam

The positive effect of universal preschool education on the matriculation rate and quality of matriculation certificate enhanced access to higher education. It is therefore likely to find an increase in the participation in the psychometric exam, which is required for university and academic college admission alongside the matriculation certificate. Indeed, as reported in the first row of Table 2, we find that the participation rate in the psychometric exam increased significantly by 2.8 percentage points (a 7% increase) when we examine whether individuals ever took this exam, and by 3.3 percentage points (a 9% increase) when we examine whether individuals took it by age 19.[16] We find an effect for both genders, with a larger impact for boys, who have a lower baseline mean than girls.

We also examine performance on the psychometric exam. Preschool education could improve test performance due to the increase in schooling, the improvement in verbal skills (Gormley and Gayer, 2005), an increase in IQ (Elango et al., 2016), and the improvement of additional soft skills associated with test performance, such as academic motivation and behavioral self-regulation (Heckman et al., 2013). To examine test performance, we construct a series of indicators for performance above different quartiles of the test score distribution to avoid selection bias from an increased probability of taking the exam.[17]

[16] We examine the outcome of taking the exam by age 19 to focus on a result that does not suffer from censoring.

[17] Students can take the psychometric exam multiple times and choose their best score for application to institutions of higher education. The table reports the results on the maximum score attained. Results using the first score are similar and available upon request.

The indicators get a value of zero for students who did not take it.[18] Estimates for the test score indicators suggest that universal preschool education improved individuals' total score as well as the score in each section: Verbal (Arabic), Quantitative, and English. We observe positive effects not only for score threshold indicators at the bottom of the test score distribution (probably induced by the increase in the number of test takers) but also for threshold indicators in the middle part of the distribution, and sometimes even for the top part of the distribution. Generally, the effect is larger for boys than for girls.

Enrollment in Postsecondary Institutions

We next examine the effects of universal preschool on enrollment in postsecondary education, an important milestone in the educational trajectory. Preschool may foster early cognitive and socio-emotional skills that shape educational aspirations and readiness for higher learning environments.

We cannot fully observe postsecondary enrollment for all cohorts, as the youngest (born in 1999) was only 18–19 in the final year of our data (2018). We therefore focus on the 1991–1998 cohorts and examine two outcomes: enrollment at any age and enrollment by age 19, the most common age of postsecondary entry among Israeli Arabs (Figure A4).[19]

The results reported in Table 3 show that preschool education had substantial effects beyond the reported increase in high school achievement. Focusing on the estimates that denote enrollment at any age (columns (1)–(3)), we see an increase in the probability of enrollment in any postsecondary education institution by 5.3 percentage points (a 16% increase over the pre-reform mean). This effect is pronounced at almost all levels of postsecondary education: first-tier university education, second-tier academic college education, and vocational education. We also see a decrease in the probability of attending teacher training institutions.[20] Note that the decline in enrollment in teacher training institutions is smaller

[18] The quartiles are defined based on the full distribution of test scores of exams in the Arabic language in 2015, which is roughly the middle of the sample period (NITE, 2017, pp. 13 and 303). Test scores in the Arabic version of the exam are much lower than in the Hebrew one. In 2015, for example, the average total score of students who took the exam in Hebrew was 576, whereas the average total score of students who took the exam in Arabic was 477. The gap in that year accounts for 0.9 of a standard deviation.

[19] The advantage of the first outcome—postsecondary enrollment at any age—is that it is more inclusive and captures individuals over a longer time horizon. However, it is subject to censoring, particularly for the younger cohorts. Nevertheless, it remains informative about the effects of universal preschool, provided that enrollment timing is similar across treatment and comparison localities and is adequately controlled for using cohort fixed effects. For robustness, we also examine an uncensored outcome: postsecondary enrollment by age 19.

[20] Teacher training institutions are the least selective postsecondary academic institutions. In the 2017/2018 academic year, the average psychometric score of students enrolled in these institutions (488) was significantly lower than that of students enrolled in universities (628) and in academic colleges (521) (CBS, 2019a, 2019b).

than the increase observed in other institutions, implying that the latter stems both from an increase in postsecondary enrollment and from some switching of individuals from teacher training institutions to institutions of higher quality. Our findings are qualitatively similar when we examine an uncensored outcome: postsecondary enrollment by age 19 (columns (4)–(6)). There are some differences by gender for the uncensored outcomes, but once we examine the effects in percentage terms (relative to the outcome means), the impact seems to be similar for boys and girls, with a slightly larger increase for boys. For example, the probability of postsecondary enrollment by age 19 increased by 24% for boys and 21% for girls.

Additional Outcomes

Juvenile Crime

Arabs are disproportionately represented in criminal activity records in Israel. In 2019, Arab youth accounted for 35% of juvenile criminal records while their share in the population was only 28% (Knesset Research and Information Center, 2020). Furthermore, in 2019, 20% of Arabs reported that they did not feel safe from violence in their locality of residence, compared to only 8% of Jews (CBS, 2021a). In our study population, the share of males with at least one criminal conviction in their juvenile record (by age 18) was 17% in the pre-reform cohorts of the treatment localities.

Preschool education might reduce juvenile crime by improving personality traits and reducing externalizing behavior (Heckman et al., 2013), and by keeping children longer in school and thus off the streets (Lochner and Moretti, 2004). It might also affect individual preferences for crime, instilling moral values and increasing the psychic costs of breaking the law (Arrow, 1997), as well as increase patience, inducing individuals to avert risky behaviors (Becker and Mulligan, 1997).

Table 4 shows that universal preschool indeed reduced the likelihood of having a juvenile crime record by 3 percentage points for boys (an 18% decrease over the pre-reform mean). The reduced rate stems from a decline in life/body offenses and in sex/property offenses.[21] Interestingly, the effect on security/order offenses is much smaller and insignificant. This is in line with the literature that finds no causal relationship between education or economic conditions and terrorism or hate crimes (see, e.g., Krueger and Malečková, 2003; Abadie, 2006; Benmelech et al., 2012). Estimates for the effects of

[21] Security/order offenses include offenses against the security of the state or against public order. Life/body offenses include offenses against a person's life and bodily harm. Sex/property offenses include sexual offenses and property offenses. Other offenses include fraud, morality offenses (usually drug-related), economic offenses, licensing offenses, and administrative offenses. Our data does not include a more detailed breakdown of the offenses for confidentiality reasons.

preschool education on juvenile crime among girls are essentially zero. This finding is expected given the low baseline mean for girls (less than 0.5% versus 17% for boys).

Early Marriage

Although Israeli Arabs went through a rapid modernization process in the last half century, they remain a relatively traditional society. In 2017, the average age of first marriage was 23 years for Israeli Arab women in contrast to an average age of 26 years for Israeli Jewish women and 30 years for women in the OECD countries.[22] Figure A5 presents the cumulative share of married men and women between the ages of 17 and 27 in the 1991 cohort (pre-treatment cohort), for which we can observe the longest time horizon. As the figure shows, a notable portion of the women, about one third, married at early ages (18–21), compared to 2% of men. We examine the effect of preschool education on marriage by age 21, since we can observe this outcome across several post-reform cohorts without censoring.

Preschool education could potentially delay the age of first marriage by reducing the probability of dropping out of high school and by increasing the probability of enrollment in higher education institutions, as documented above. In addition, the better employment and earnings prospects of educated women are expected to reduce gains from marriage in a framework where men and women specialize in market and non-market work, respectively, as is typical of traditional societies (Becker, 1981; Blau et al., 2000). Finally, increased education might affect the age of marriage by reducing religiosity and eroding traditional values (Hungerman, 2014; Cesur and Mocan, 2018).

The effects of universal preschool on the probability of marrying at an early age are presented in Figure 4, where we plot DID estimates and 95% confidence intervals from models in which the dependent variable is marrying by age 18, 19, 20, or 21. Panel A reports estimates for women. The estimates are a bit noisy, but they all point to a decline of about 1.5–2 percentage points in the probability of early marriage. Regarding marriage by age 21, the point estimate implies a decline of 5% relative to a baseline of 32%. Panel B reports estimates for men, which are small, with confidence intervals that do not reject the hypothesis of a zero effect.[23]

[22] The statistics for Jews and Arabs were calculated by the authors from Tables 2.35 and 2.36 in CBS (2020). OECD statistics are taken from Indicator SF3.1 in OECD (2019).

[23] Estimates for marriage of males by age 18 are not included since there are almost no married males by this age in the sample.

6. Robustness, Falsification Tests, and Family Fixed Effects

Robustness and Falsification Tests

We conduct several robustness tests to assess the feasibility of our identification assumption and ensure that our findings are not driven by unobserved differential trends in the treatment and comparison localities or specific shocks that differentially affected younger versus older cohorts in treated and comparison localities. We describe and report these tests in detail in Appendix D and summarize them here.

We first assess the sensitivity of our results to the inclusion of the set of background characteristics used in our main specification. We also estimate models that include a linear time trend interacted with a locality's socioeconomic cluster or socioeconomic ranking (together with the baseline linear trend) (Table A5). All estimates are similar to our main results. Moreover, we observe no major differential change in student background characteristics between the pre- and post-cohorts of treatment and comparison localities (Table A6), suggesting our results are not driven by demographic or compositional changes in the localities. We also conduct a placebo analysis where we estimate our main DID model using only the pre-reform cohorts, assuming the law was implemented two years before it actually was (Table A7). These tests show no evidence of significant differential pre-reform trends between treatment and comparison localities, supporting our main identification assumption.

We also examine the robustness of our results for different dynamics in outcomes in different type of localities or ethnic groups (Table A8). In particular, we show that our estimates are similar when we include in the comparison group only the always-treated or only the never-treated localities or when we exclude different subgroups of the treatment or comparison sample (e.g., Bedouin or Druze). We also show that our results are not driven by any specific locality by re-estimating our model dropping one locality each time (Figure A6). Moreover, we address potential small sample bias from having only 37 clusters (15 treated) by implementing wild bootstrap procedures, which produced p-values nearly identical to those obtained using the standard cluster adjustments (Table A9).

To address concerns that the estimated effects are driven by changes in inputs at later stages of schooling, we examine whether there were differential changes in class size in elementary and secondary school between treated and comparison localities, finding no evidence of such differences (Table A10). Similarly, we find no evidence for differential changes in per capita expenditure, education expenditure per capita (ages 0-17), or revenue per capita between treatment and comparison localities after 1999 (Table A11). 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.[24]

Last, to reduce the concerns that our results are driven by a specific shock in 1999/2000 that differentially affected young versus older children in treatment versus comparison localities, we add five localities to our sample that were treated after 2000 due to reclassification of their socioeconomic status and re-estimate equation (1) for high school outcomes. Results remain highly similar, supporting our causal interpretation of the findings (Table A12).

Family Fixed Effects

Our comprehensive data allow us to identify siblings and estimate a model with family fixed effects. We compare the outcomes of children young enough to have access to universal preschool to their older siblings who were already over the age of 4 when the reform was implemented in treatment localities and the outcomes of children and siblings born in the same years in comparison localities. The estimated equation is similar to equation (1), replacing the locality fixed effects with a family fixed effect based on the mother's individual identifier.

Comparison of the estimates from the family fixed effects model with the estimates from the main DID model also provides insights into the extent of intra-household resource allocation. For example, a larger impact within rather than across families would suggest that parents reinforce differences in human capital investments between their children. By contrast, a smaller impact within rather than across families would indicate compensatory behavior, where parents attempt to reduce human capital gaps

[24] 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.

among their children. Alternatively, differences between estimates from the two models might point to unobserved trends or shocks at the locality level that could have biased our baseline DID estimates.

Table 5 reports the estimates of the family fixed effects model (column (3)). To ease comparison, we report the estimates of the baseline DID model in column (1) and DID model estimates after restricting the sample to families with at least two children (matching the sample of the family fixed effects model) in column (2). The estimates of DID model based on the restricted sample are almost identical to our main estimates, although they are slightly less precise because of reduced sample size. The estimates of the family fixed effects model are remarkably similar to those of the DID model but are slightly noisier due to the addition of family fixed effects. The similarity of the estimates in the two models provides further evidence of the validity of our main identifying assumption, minimizing concerns that our results are driven by time-varying shocks that differentially affected households in treated versus comparison localities.[25]

7. Heterogeneity Analysis, Mechanisms, and Intermediate Outcomes

Early childhood interventions are generally more beneficial among disadvantaged populations (Blau and Currie, 2006; Elango et al., 2016). One critical factor when examining heterogeneity of preschool programs is the counterfactual childcare. This is particularly important in the case of universal preschool provision as it might crowd out high-quality targeted programs (e.g., Bassok et al., 2014). Alternatively, universal preschool might provide an educational environment for children who would have otherwise been at home or attended low-quality childcare. Evidence on at-home care versus formal childcare points to beneficial effects for children from lower SES families (Cascio and Schazenbach, 2013; Felfe et al., 2015; Drange and Havnes, 2019) and mixed or detrimental effects for children from high SES families (Herbst, 2013; Havnes and Mogstad, 2015). In our setup, the counterfactual childcare was mainly home care by the mother or a close relative. Therefore, the results should be interpreted in this context.

Another important issue when analyzing heterogeneity across groups is the compliance rate for each group. Unfortunately, we lack data on preschool enrollment at the individual level for the pre-reform period.[26] Nevertheless, to gain some insights on the characteristics of compliers, we examine differences

---

[25] Family fixed effects models identify impacts for "switcher" families (those with treated and untreated children), which may differ from the main sample (Miller et al., 2023). In our case, the proportion of such families is relatively large. Moreover, their background characteristics and mean outcomes are very similar to the main sample (see Table A13). Thus, estimates obtained from family fixed effects models and our main specification are based on similar samples. We discuss this further in Appendix E.

[26] Data on preschool enrollment in the pre-reform period is only available at the aggregate locality level.

in preschool attendance by family background in the post-reform period between treated localities and never-treated localities in the comparison group (Table A14). Overall, the analysis shows no consistent pattern of selection into preschool enrollment by sociodemographic characteristics, implying that the universal preschool policy reached children from all socioeconomic backgrounds.[27] These results imply that our ITT estimates among different groups (reported below) reflect differences in the impact of preschool attendance rather than differences in compliance.

Table 6 displays DID estimates and outcome means for the effects of universal preschool for different groups, focusing on the summary outcomes from our main analysis for each domain. Our findings for other outcomes align consistently with the results discussed below.[28] Given the extremely low incidence of juvenile crime among girls and of early marriage among boys, we report estimates for the relevant genders for these two outcomes (crime for boys and marriage for girls); for all other outcomes, we use the full sample.

Estimates obtained from the stratification by parental education (columns (1)–(4)) suggest that the positive effects of universal preschool human capital outcomes are stronger among children whose parents, especially mothers, did not complete 12 years of schooling. We do not find heterogeneity by parental education for social outcomes (juvenile crime and early marriage).[29]

We also examine heterogeneous effects for father's income (columns (5) and (6)) and mother's employment (columns (7) and (8)), both measured when children were two years old. For the analysis by father's income, we stratified the sample by real annual income below or above the sample median (28,400 NIS, equivalent to $8,200 in 2021).[30] The impact of universal preschool tends to be similar for children from low- versus high-income fathers for most outcomes, while the decline in women's early

[27] Estimates from a linear probability model of preschool attendance at age 3 and 4 are reported in columns (1) and (2), respectively. The model includes background characteristics and their interactions with a treatment dummy. Overall, most estimates show no significant pattern of selection, except one estimate of positive selection into enrollment at age 3 by father's education and one estimate of negative selection into enrollment at age 4 by number of siblings.

[28] We also estimated heterogeneous effects stratifying the sample by number of siblings and by parity (first-born versus later-born children). Estimates (not reported here to save space) did not point to a consistent pattern of heterogeneity.

[29] Several factors may explain the similar crime-reducing effects of preschool education across maternal education groups despite the different effect for academic outcomes. First, the smaller baseline differences between groups in criminal activity compared to academic outcomes suggest that factors beyond family background may have a stronger influence on criminal behavior. Second, preschool may boost soft (non-cognitive) skills broadly for all groups. Last, different mechanisms (e.g., "incarceration effects" for less-educated families; increased opportunity costs for more-educated families) may produce similar crime reductions. The examination of these different mechanisms is left for future research.

[30] We assign a value of zero to fathers with no earnings during the year. Therefore, the annual median income is quite low.

marriage comes mainly from households with low-income fathers. The impact of preschool on human capital outcomes is higher for children of non-employed mothers.

We also examine heterogeneity in treatment effects with respect to children's predicted outcomes computed based on models that use the pre-treatment cohorts and students' background characteristics.[31] For each outcome of interest, we divide the entire population into tertiles based on the value of the predicted outcome and estimate equation (1) separately for each tertile. This allows us to study effect varies across individuals whose expected performance would have been low, medium, or high absent the reform. Results appear in Table 7.[32] The effects on high school performance is significantly larger for those with low and medium predicted outcomes than for those located in the highest tertile. Notably, for the latter group we see a substantial increase in postsecondary enrollment.[33] Our results are similar when we stratify the sample using a single predicted outcome: the likelihood of obtaining a matriculation certificate, and estimate our DID model for all outcomes based on this stratification (Table A16). Overall, universal preschool education benefited different children at different margins. It had a large impact on high school performance among the most disadvantaged and benefited more advantaged children by improving post-secondary enrollment.[34] Our results stress the importance of studying multiple outcomes across different population groups to properly assess the effects of universal preschool education.

---

[31] We use the pre-treatment cohorts to estimate separate OLS or logistic models for boys and girls, regressing each outcome on the set of student background characteristics specified in equation (1). Using these estimated coefficients, we predict outcomes for each student in our sample and then classify students into tertiles (low/medium/high) by their predicted outcomes.

[32] Table A15 displays estimates from a fully saturated model, used to assess which differences in treatment effects between low/medium predicted outcomes and high predicted outcomes are statistically significant.

[33] We also find a slightly higher decrease in crime among individuals with a low predicted probability of arrest (more advantaged individuals), but the difference is not statistically significant (Table A15).

[34] There may be different mechanisms through which children from relatively more advantaged backgrounds and higher academic potential benefit from preschool education. First, they might have developed cognitive and non-cognitive skills that are particularly relevant for higher education. Indeed, we found improved psychometric exam scores among children with higher predicted outcomes (results not reported but available upon request). Second, even if the improvement in high school achievement was more modest for children with higher predicted outcomes, we do find a positive effect on the index of high school performance, suggesting that universal preschool also affected their readiness for higher education. Third, children from more advantaged backgrounds might have also benefited indirectly from the improved learning environment and the higher instructional level of classes enabled by the equalizing stage of preschool education for all children.

Intermediate Outcomes in Elementary and Middle School

Test Scores

To investigate potential mechanisms for the effects found on individuals' long-term outcomes, we also examine intermediate outcomes in elementary and middle school. We focus on a subsample of individuals for which we have data on achievement in the GEMS exams (*Meitzav*) in fifth and eighth grades. As not all schools are tested every year, 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.

Figure 5 displays estimates of this DID specification with 95% confidence intervals on the effects of universal preschool on z-scores. Given the smaller sample size, our results are less precise. The most pronounced effect of universal preschool is on individuals' native language skills (Arabic), with scores increasing significantly by 0.11 standard deviations in fifth grade. Notably, the effect persisted in eighth grade, where scores improved by 0.17 standard deviations. We also find an effect on math scores of 0.20 standard deviations in fifth grade, but no such effect in eighth grade. Thus, either the beneficial effects on math achievements diminish over time (as in Deming, 2009, and other studies examining short- versus long-term effects of preschool education), or the math skills tested in fifth grade are not highly correlated with those tested in eighth grade. Our results are consistent with Felfe et al. (2015), who examined the effects of a universal preschool reform in Spain during the 1990s on tenth-grade achievement scores, finding a 0.15 increase in reading scores but no effect on math achievements. The large improvement in Arabic test scores may explain the sharp increase in enrollment in higher education documented in Section 5. This aligns with the findings of Aucejo and James (2021), who assert that verbal skills play a pivotal role in explaining the variation in university enrollment among individuals, with their marginal effect being more than twice as large as that of math skills.

We find no significant effect of universal preschool education on children's performance in English and science in the fifth and eighth grades. While this ostensibly contradicts some of our main findings (a significant increase in the number of English units and science subjects tested in the high school matriculation exams), note that science and English skills are not directly taught in preschools. It is more likely that participation in preschool boosted children's non-cognitive skills, such as academic motivation, persistence, and learning initiative, needed to succeed in the matriculation exams (see Heckman et al., 2013). This explanation is further supported by the distinction between matriculation exams, which are high-stake tests influencing access to higher education and certain jobs, and the GEMS tests, which are low-stake assessments designed to evaluate general trends in Israel's public education system.[35]

Learning Environment

The improvement in achievement of students exposed to preschool could derive not only from greater cognitive and non-cognitive skills, but also from a better learning environment. Preschool develops crucial soft and socio-emotional skills that influence behavior, self-regulation, and self-discipline throughout a child's education (e.g., Heckman et al., 2013). These enhanced skills are likely to be reflected in an improved learning environment and classroom climate. We use data from the GEMS questionnaire administered to students in grades 5–9 to examine how universal preschool education affected the learning environment. Though we lack data on the learning environment in high school, we expect similarities, as there is substantial continuity in learning environments from elementary and middle school through high school in our study context.[36] Moreover, an improved learning environment in elementary and middle school could help students develop stronger academic foundations, study habits, and peer relationships that support their educational trajectory and ultimately translate into higher achievement in high school.

Our specification is similar to equation (1), where we control for the type of school (Druze, Bedouin, or other Arab) and fixed effects for cohort, locality, grade, and test year. We do not include student covariates, as the questionnaires are completely anonymized. All outcomes are defined as binary outcomes that take the value of 1 if the student strongly to partially agrees with a given statement and 0 if the student strongly to partially disagrees. We also construct a z-score index that aggregates all dimensions of the school environment. This index is computed as the average of all standardized individual environment variables, similar to the previously estimated z-score of high school performance.

[35] Differences in results across school subjects are not driven by ceiling effects or insufficient variation in test scores, as the standard deviation is similar across subjects and grades. A more plausible explanation is that some subjects are more strongly related to long-term outcomes. Scores in fifth and eighth grade for all four subjects predict long-term outcomes even after adjusting for background characteristics, with Arabic and math scores having stronger predictive power for matriculation and postsecondary enrollment, consistent with the larger treatment effects observed for these subjects.

[36] The learning environments observed in elementary and middle school strongly predict those in high school due to institutional features of the education system. In the treated localities, most students not only share the same cohort of peers throughout their educational trajectory, but also often attend combined middle-high school institutions. Our data show that, for the average student, approximately 70% of their tenth-grade classmates were also their eighth-grade peers (for the median student, this value is 90%). Such continuity means that improvements in the learning environment at earlier stages, especially middle school, likely persist into high school.

Table 8 shows that students who received universal preschool education experienced a better learning environment in elementary and middle school, as they were significantly more likely to report that they enjoyed school (4.2 percentage points, a 6% increase) and that students tended to help each other in class (2.1 percentage points, a 2.7% increase). They were also significantly less likely to report frequent classroom disturbances (4.2 percentage points, a 6% decrease).

Students in the treated cohorts also reported a greater sense of safety and security, as well as better teacher–student relationships. They were 5.3 percentage points (18%) less likely to report that they are sometimes afraid to attend school because there are violent students, and they were 3 percentage points (4%) more likely to report that teachers help prevent violence and maintain discipline. Moreover, the percentage of students who reported having a good relationship with teachers increased by 2.5 points (3%), and students were 7.2 percentage points (16%) less likely to report being insulted by teachers.

To rule out the possibility that these findings stemmed from unobserved differential trends or confounding factors, we also examined effects on additional student questionnaire items not expected to be affected by universal preschool education, such as computer use at home and at school across different subjects. Reassuringly, the estimated effects for all these outcomes (reported in the right column of the table) are insignificant. The absence of effects on school computer use also suggests that the positive effects on educational outcomes are unlikely to be confounded by an increase in school inputs in treated localities for the cohorts that received universal preschool education.

In sum, one possible mechanism explaining the impact of universal preschool on long-term outcomes is the creation of a safer and more conducive learning environment. These findings suggest that the provision of universal preschool affected not only the *complier* population of children who enrolled in preschool as a result of the law, but also the entire cohort of students and their teachers in treatment localities. All benefited from the enhanced learning environment.[37]

Maternal Employment

One possible channel that could explain the improved outcomes of children with access to universal preschool is an increase in maternal employment and thus household income. In traditional societies such as the Arab community, women were the primary caregivers for children. Access to universal preschool

[37] The improved learning environment among treated cohorts likely stems from both individual behavioral improvements and peer effects. While we cannot fully disentangle these channels, they are inherently connected as treated students typically attend schools with treated peers. Importantly, we can rule out alternative explanations for the improved learning environment. As detailed in Table A17, universal preschool did not simply redirect students to higher quality high schools as measured by pre-treatment matriculation rates.

due to the Preschool Law might have encouraged them to work and increased household income. However, we find no evidence of a significant increase in mothers' employment or income during the study period, ruling out these channels of impact (see Appendix F and Tables A17 and A18).

8. Comparison with Other Preschool Programs and with Alternative School Interventions in Israel

To put our results in perspective, we compare them to results obtained in the literature for other universal or large-scale preschool education programs, as well as small-scale targeted programs. So far, we have reported intention-to-treat (ITT) estimates for the effects of universal preschool education, which are interesting for policy purposes as they shed light on the effect of providing access to universal preschool education. They also provide information on the overall effect of universal preschool education on all children, including those who did not attend public preschool but lived in treatment localities and could have been indirectly affected. To compare our results with those of other studies, we report here local average treatment effects (LATE) by scaling up our intention-to-treat (ITT) estimates by the increase in public preschool enrollment generated by the reform (about 60 percentage points).[38]

Table 9 compares our estimates with those from other preschool education studies. For proper context, we report (where available) baseline means, counterfactual care arrangements, and maternal education levels for each study. With the exception of Gray-Lobe et al. (2023), children in other large-scale programs would have faced similar counterfactual care arrangements and come from families with relatively low levels of education, as in our study. Nevertheless, given the diverse contexts, populations, and educational systems across these studies, we do not necessarily expect our results to align with findings from other settings.

We focus on the most comparable outcomes across studies: high school graduation and college enrollment. The ITT effect on high school graduation in our study is 0.028, implying a LATE estimate of about 5 percentage points (a 6% increase relative to the baseline outcome mean). This is within the range of other studies examining the effects of large-scale preschool education programs (Table 9, Panel A). Note, however, that it is at the lower end of the distribution, which might be explained by the higher baseline mean for our study population. In fact, there seems to be a negative relationship between the effect of preschool education on high school graduation rates and the baseline outcome mean when we compare across studies. In contrast, we observe a much larger effect on college enrollment in our study

[38] Table A20 reports DID estimates for the effects on public preschool enrollment based on aggregate data at the locality level weighted by population size.

than in other large-scale studies: 6.7 percentage points (26%). This might derive from a lower baseline college enrollment rate in our sample population than in other studies.

Panel B of Table 9 summarizes results from the literature on targeted programs. Our estimates are smaller for both outcomes than those obtained in targeted programs. Nevertheless, most of these studies find beneficial effects largely for girls, while we find that universal preschool education increased human capital for both genders.

In Table 10, we compare our results with estimates from studies examining the impact of educational interventions in Israel targeted at older children during the same period. We focus on two high school interventions that report causal estimates for a subset of comparable outcomes, comparing the costs of each intervention and the estimated gains.[39] Lavy and Schlosser (2005) examine the effects of remedial education for underperforming high school students who were at the margin of obtaining a matriculation certificate. The per-student cost of this intervention was $1,100, compared to an estimated cost of $1,400 for universal preschool provision. Remedial education generated an increase of 13 percentage points in the probability of obtaining a matriculation certificate among treated students. The effect in absolute terms is larger than that of universal preschool education (13 versus 7 percentage points), and improvement over the outcome means is 24% for remedial education and 17% for universal preschool education. Nevertheless, the effect of universal preschool education is substantially larger in the long term: Lavy et al. (2022) find an 8-percentage point increase (15%) in enrollment to low-tier higher education institutions (colleges and vocational institutions), with no effect on enrollment in high-tier institutions (universities). In our study, universal preschool education increased enrollment in higher education institutions by 9 percentage points (27%), with positive effects in almost all tiers, including universities.

The second intervention, examined by Angrist and Lavy (2009), provided monetary awards to students in low-achieving high schools based on their success in the matriculation exams. The cost of the intervention was relatively low, only $385 per student, as the award was only given to students who achieved the target. The authors find a significant increase of 14 percentage points in the probability of obtaining a matriculation certificate for girls, with no significant effect for boys. Although this is a larger effect on matriculation rates than in our study, they find no effect in the longer term on university

[39] The two interventions were implemented during the same period on different cohorts, so there is no concern of overlap between populations. Moreover, only a small proportion of Arab students participated in these interventions. Unfortunately, these studies do not report separate estimates for the Arab population, as the subsample of Arab students is relatively small in each.

enrollment and only a localized effect on postsecondary enrollment in second-tier institutions for girls in the top quartile of the achievement distribution.

In sum, universal preschool education entails higher costs than the two high school interventions, but the longer-term benefits in postsecondary enrollment appear to be significantly larger. A more comprehensive comparison should include the rate of return in terms of dollars spent and embed the monetary benefits of additional outcomes, such as criminal activity, early marriage, and fertility rates. We plan to assess this in future work, when the cohorts exposed to universal preschool education enter the labor market.

9. Summary and Conclusions

This study provides a comprehensive set of findings on the impact of universal preschool education within a disadvantaged population: the Arab population in Israel. Our results indicate that access to universal preschool at ages 3 and 4 benefited individuals across various stages. It enhanced their language skills in elementary and middle school and raised their performance in fifth-grade math exams. In high school, universal preschool education decreased the likelihood of dropping out, raised participation in the matriculation exams, increased eligibility for a matriculation certificate, and improved the quality of the certificate achieved (reflected in the number of math and English units and the number of science subjects). The probability of enrollment in postsecondary education also increased significantly for both academic and vocational institutions. We also find beneficial effects of universal preschool education on other long-term outcomes: a decline in the probability of engaging in juvenile crime among boys and in the probability of marrying at an early age among girls.

Possible mediating factors for the long-term benefits of universal preschool education include significant improvements in the learning environment. Students reported greater enjoyment of school, a higher sense of safety, fewer in-class disturbances, and better enforcement of discipline in the classroom, as well as better relationships with teachers and classmates.

We find that universal preschool education affected different children at different margins. It had a larger impact on high school performance for children from low or medium socioeconomic backgrounds, whereas it increased the probability of postsecondary enrollment for children from higher socioeconomic backgrounds. The long-term impact of universal preschool education on postsecondary enrollment is larger than in other educational interventions implemented in Israel among high school students during the same period, emphasizing the importance of human capital investments at younger ages.

One possible lesson from our study is that disadvantaged communities can benefit from universal preschool education, even in the absence of well-targeted educational programs. Free universal preschool education can provide stimuli, enhance social experience, and enrich verbal skills for disadvantaged children, which they cannot always get in their family environment. While there is a growing interest in the effects of universal preschool education on individuals' outcomes and achievements, there are almost no studies that examine its implementation in a traditional non-Western society. We believe that the Arab-Israeli experience can be a useful example, showing positive short- and long-term benefits of universal preschool education for disadvantaged communities.

References

Abadie, A. (2006). Poverty, political freedom, and the roots of terrorism. *American Economic Review, 96*, 50–56.

Abu Ahmad, H., Ibrahim, R., & Share, D. L. (2014). Cognitive predictors of early reading ability in Arabic: A longitudinal study from kindergarten to grade 2. In E. Saiegh-Haddad, & R. M. Joshi (Eds.), *Handbook of Arabic Literacy: Insights and Perspectives* (Vol. 9, pp. 171–194). Dordrecht: Springer.

Abu-Bader, S., & Gottlieb, D. (2013). Poverty, education, and employment among the Arab-Bedouin in Israel. In *Poverty and social exclusion around the Mediterranean Sea* (pp. 213–245). Boston: Springer.

Abu-Jaber, G. (1994). *Early childhood education in the Arab sector: Report from a field survey in January-July 1993.* Jerusalem: Shatil.

Abu-Rabia, S. (2000). Effects of exposure to literary Arabic on reading comprehension in a diglossic situation. *Reading and Writing, 13*, 147–157.

Anderson, M. L. (2008). Multiple inference and gender differences in the effects of early intervention: A reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects. *Journal of the American Statistical Association, 103*, 1481–1495.

Angrist, J., & Lavy, V. (2009). The effects of high stakes high school achievement awards: Evidence from a randomized trial. *American Economic Review, 99*, 1384–1414.

Aram, D., Korat, O., & Hassunah-Arafat, S. (2013). The contribution of early home literacy activities to first grade reading and writing achievements in Arabic. *Reading and Writing, 26*, 1517–1536.

Arrow, K. J. (1997). The benefits of education and the formation of preferences. In J. R. Behrman, & N. G. Stacey (Eds.), *The Social Benefits of Education* (pp. 11–16). Ann Arbor: University of Michigan Press. doi:10.3998/mpub.15129

Aucejo, E., & James, J. (2021). The path to college education: The role of math and v Skills. *Journal of Political Economy, 129*, 2905–2946.

Bailey, M. J., Sun, S., & Timpe, B. (2021). Prep school for poor kids: The long-run impacts of Head Start on human capital and economic self-sufficiency. *American Economic Review, 111*, 3963–4001.

Bank of Israel. (2002). *Recent economic developments 99, April-September 2002.* Jerusalem.

Bank of Israel. (2003). *Recent economic developments 100, July-December 2002.* Jerusalem.

Bank of Israel. (2021). Welfare Policy Issues. In *Annual Report 2020* (pp. 224–261). Jerusalem: Bank of Israel.

Bassok, D., Fitzpatrick, M., & Loeb, S. (2014). Does state preschool crowd-out private provision? The impact of universal preschool on the childcare sector in Oklahoma and Georgia. *Journal of Urban Economics, 83*, 18–33.

Becker, G. S. (1981). *A treatise on the family.* Cambridge, MA: Harvard University Press.

Becker, G. S., & Mulligan, C. B. (1997). The endogenous determination of time preference. *The Quarterly Journal of Economics, 112*, 729–758.

Belfield, C. R., Nores, M., Barnett, S., & Schweinhart, L. (2006). The High/Scope Perry Preschool program cost–benefit analysis using data from the age-40 followup. *Journal of Human Resources, 41*, 162–190.

Benmelech, E., Berrebi, C., & Klor, E. F. (2012). Economic conditions and the quality of suicide terrorism. *The Journal of Politics, 74*, 113–128.

Berlinski, S., Galiani, S., & Gertler, P. (2009). The effect of pre-primary education on primary school performance. *Journal of Public Economics, 93*, 219–234.

Berlinski, S., Galiani, S., & Manacorda, M. (2008). Giving children a better start: Preschool attendance and school-age profiles. *Journal of Public Economics, 92*, 1416–1440.

Blanden, J., Del Bono, E., McNally, S., & Rfabe, B. (2016). Universal pref-school education: The case of public funding with private provision. *The Economic Journal, 126*, 682–723.

Blass, N., & Adler, C. (2004). Politics, education and scientific knowledge – Is there any connection? *Megamot, 1*, 10–32.

Blau, D., & Currie, J. (2006). Pre-school, day care, and after-school care: Who's minding the kids? *Handbook of the Economics of Education, 2*, 1163–1278.

Blau, F. D., Kahn, L. M., & Waldfogel, J. (2000). Understanding young women's marriage decisions: The role of labor and marriage market conditions. *ILR Review, 53*, 624–647.

Brown, M. C., Sibley, D. E., Washington, J. A., Rogers, T. T., Edwards, J. R., MacDonald, M. C., & Seidenberg, M. S. (2015). Impact of dialect use on a basic component of learning to read. *Frontiers in Psychology, 6*, 196.

Bruhn, J., & Emick, E. (2023). *Lottery evidence on the impact of preschool in the United States: A review and meta-analysis.* Technical Report, Blueprint Labs.

Callaway, B., & Sant'Anna, P. H. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics, 225*, 200–230.

Cameron, A. C., Gelbach, J. B., & Miller, D. L. (2008). Bootstrap-based improvements for inference with clustered errors. *The Review of Economics and Statistics, 90*, 414–427.

Campbell, F. A., Pungello, E. P., Burchinal, M., Kainz, K., Pan, Y., Wasik, B. H., . . . Ramey, C. T. (2012). Adult outcomes as a function of an early childhood educational program: An Abecedarian Project follow-up. *Developmental Psychology, 48*, 1033–1043.

Cascio, E. U. (2023). Does universal preschool hit the target? Program access and preschool impacts. *Journal of Human Resources, 58*, 1–42.

Cascio, E. U., & Schanzenbach, D. W. (2013). *The impacts of expanding access to high-quality preschool education.* National Bureau of Economic Research.

CBS. (2000). *Statistical abstract of Israel No. 51.* Jerusalem: Central Bureau of Statistics.

CBS. (2003). *Characterization of geographic units and their classification - according to the socio-economic level of the population in 1999.* Jerusalem: Central Bureau of Statistics.

CBS. (2019a). *Applications to first degree studies at universities and academic colleges. press release 102/2019.* Jerusalem: Central Bureau of Statistics.

CBS. (2019b). *Trends in teacher training, specialization in teaching and entering the field of teaching, 2000-2019. Press Release 184/2019.* Jerusalem: Central Bureau of Statistics.

CBS. (2020). *Statistical Abstract of Israel No. 71.* Central Bureau of Statistics.

CBS. (2021a). Sense of personal security – Findings from the Personal Security Survey. Press Release 10/2021. Jerusalem: Central Bureau of Statistics.

CBS. (2021b). *Statistical Abstract of Israel No. 72.* Central Bureau of Statistics.

Cesur, R., & Mocan, N. (2018). Education, religion, and voter preference in a Muslim country. *Journal of Population Economics, 31*, 1–44.

Cornelissen, T., Dustmann, C., Raute, A., & Schönberg, U. (2018). Who benefits from universal child care? Estimating marginal returns to early child care attendance. *Journal of Political Economy, 126*, 2356–2409.

Cunha, F., & Heckman, J. (2007). The technology of skill formation. *American Economic Review, 97*, 31–47.

Currie, J., & Almond, D. (2011). Human capital development before age five. In *Handbook of labor economics* (Vol. 4, pp. 1315–1486). Amsterdam: Elsevier.

Deming, D. (2009). Early childhood intervention and life-cycle skill development: Evidence from Head Start. *American Economic Journal: Applied Economics, 1*, 111–34.

Drange, N., & Havnes, T. (2019). Early childcare and cognitive development: Evidence from an assignment lottery. *Journal of Labor Economics, 37*, 581–620.

Duncan, G., Kalil, A., Mogstad, M., & Rege, M. (2023). Investing in early childhood development in preschool and at home. *Handbook of the Economics of Education, 6*, 1–91.

Elango, S., García, J. L., Heckman, J. J., & Hojman, A. (2016). Early childhood education. In *Economics of means-tested transfer programs in the United States* (Vol. 2, pp. 235–297). Chicago: University of Chicago Press.

Felfe, C., Nollenberger, N., & Rodríguez-Planas, N. (2015). Can't buy mommy's love? Universal childcare and children's long-term cognitive development. *Journal of Population Economics, 28*, 393–422.

Ferguson, C. A. (1959). Diglossia. *Word, 15*, 325–340.

Fitzpatrick, M. D. (2010). Preschoolers enrolled and mothers at work? The effects of universal prekindergarten. *Journal of Labor Economics, 28*, 51–85.

García, J. L., Heckman, J. J., & Ronda, V. (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*, 1477–1506.

García, J. L., Heckman, J. J., & Ziff, A. L. (2018). Gender differences in the benefits of an influential early childhood program. *European Economic Review, 109*, 9–22.

Ghanem, A. (1993). *The Arabs in Israel: Towards the 21st century, a survey of basic infrastructure.* Givat Haviva: The institute of peace research.

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics (Elsevier, 225*, 254–277.

Gormley, W. T., & Gayer, T. (2005). Promoting school readiness in Oklahoma: An evaluation of Tulsa's pre-k program. *Journal of Human Resources, 40*, 533–558.

Gray-Lobe, G., Pathak, P. A., & Walters, C. R. (2023). The long-term effects of universal preschool in Boston. *The Quarterly Journal of Economics, 138*, 363–411.

Gregory, L., Taha Thomure, H., Kazem, A., Boni, A., Elsayed, M. A., & Taibah, N. (2021). *Advancing Arabic language teaching and learning: A path to reducing learning poverty in the Middle East and North Africa.* Washington, D.C.: World Bank.

Havnes, T., & Mogstad, M. (2011). No child left behind: Subsidized child care and children's long-run outcomes. *American Economic Journal: Economic Policy, 3*, 97–129.

Havnes, T., & Mogstad, M. (2015). Is universal child care leveling the playing field? *Journal of Public Economics, 127*, 100–114.

Heckman, J. J., Moon, S. H., Pinto, R., Savelyev, P. A., & Yavitz, A. (2010). The rate of return to the High/Scope Perry Preschool program. *Journal of Public Economics, 94*, 114–128.

Heckman, J., & Masterov, D. V. (2007). *The productivity argument for investing in young children.* National Bureau of Economic Research.

Heckman, J., Pinto, R., & Savelyev, P. (2013). Understanding the mechanisms through which an influential early childhood program boosted adult outcomes. *American Economic Review, 103*, 2052–2086.

Herbst, C. M. (2013). The impact of non-parental child care on child development: Evidence from the summer participation "dip". *Journal of Public Economics, 105*, 86–105.

Hleihel, A. (2011). Barriers to internal migration among Israeli Arabs. In *Arab society in Israel: Population, society, economy* (Vol. 4, pp. 63–80). Jerusalem: Van Leer Institute and Hakibbutz Hamehuchad.

Hungerman, D. M. (2014). The effect of education on religion: Evidence from compulsory schooling Laws. *Journal of Economic Behavior & Organization, 104*, 52–63.

Israel Democracy Institute. (2022). *The inaugural annual statistical report on Arab society in Israel, 2020.* Jerusalem.

Israel State Comptroller. (1992). *State comptroller's report for 1991, No. 42.* Jerusalem: Jerusalem.

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

Kline, P., & Walters, C. R. (2016). Evaluating public programs with close substitutes: The case of Head Start. *The Quarterly Journal of Economics, 131*, 1795–1848.

Knesset Research and Information Center. (2020). *Background document for a discussion on crime and violence among youth in the Arab society.* Jerusalem.

Kop, Y. (2002). *The 2002 annual report on Israel's social services.* Jerusalem: Taub Center for Social Policy Studies in Israel.

Krueger, A. B., & Malečková, J. (2003). Education, poverty and terrorism: Is there a causal connection? *Journal of Economic Perspectives, 17*, 119–144.

Lavy, V., & Schlosser, A. (2005). Targeted remedial education for underperforming teenagers: Costs and benefits. *Journal of Labor Economics, 23*, 839–874.

Lavy, V., Kott, A., & Rachkovski, G. (2022). Does remedial education in late childhood pay off after all? Long-run consequences for university schooling, labor market outcomes, and intergenerational mobility. *Journal of Labor Economics, 40*, 239–282.

Lochner, L., & Moretti, E. (2004). The effect of education on crime: Evidence from prison inmates, arrests, and self-reports. *American Economic Review, 94*, 155–189.

Meer, J., & West, J. (2016). Effects of the minimum wage on employment dynamics. *Journal of Human Resources, 51*, 500–522.

Miller, D. L., Shenhav, N., & Grosz, M. (2023). Selection into identification in fixed effects models, with application to Head Start. *Journal of Human Resources, 58*, 1523–1566.

NITE. (2017). *Psychometric entrance exam to universities — 2015 statistical report.* Jerusalem: National Institute for Testing & Evaluation.

OECD. (2019). *OECD family database.* Paris: OECD Publishing. Retrieved from https://www.oecd.org/els/family/database.htm

Rambachan, A., & Roth, J. (2023). A more credible approach to parallel trends. *Review of Economic Studies*, rdad018.

Roth, J., Sant'Anna, P. H., Bilinski, A., & Poe, J. (2023). What's trending in difference-in-differences? A synthesis of the recent econometrics literature. *Journal of Econometrics, 235*, 2218-2244.

Saiegh-Haddad, E. (2003). Linguistic distance and initial reading acquisition: The case of Arabic diglossia. *Applied Psycholinguistics, 24*, 431–451.

Saiegh-Haddad, E. (2022). A psycholinguistic-developmental approach to the study of reading in Arabic diglossia: Assumptions, methods, findings and educational implications. In *Handbook of literacy in diglossia and in dialectal contexts: Psycholinguistic, neurolinguistic, and educational perspectives* (pp. 135–163). Cham: Springer.

Saiegh-Haddad, E., & Spolsky, B. (2014). Acquiring literacy in a diglossic context: Problems and prospects. *Handbook of Arabic literacy: Insights and perspectives*, 225–240.

Saiegh-Haddad, E., Laks, L., & McBride, C. (2022). *Handbook of literacy in diglossia and in dialectal contexts.* Cham: Springer.

Schweinhart, L., Montie, J., Xiang, Z., Barnett, W. S., Belfield, C. R., & Nores, M. (2005). *The High/Scope Perry Preschool study through age 40.* Ypsilanti MI: High/Scope Press.

Terry, J. M., Thomas, E. R., Jackson, S. C., & Hirotani, M. (2022). African American English speaking 2nd graders, verbal–s, and educational achievement: Event related potential and math study findings. *PLOS One, 17*, e0273926.

van Huizen, T., & Plantenga, J. (2018). Do children benefit from universal early childhood education and care? A meta-analysis of evidence from natural experiments. *Economics of Education Review, 66*, 206–222.

Figure 1: Geographical Distribution of the Localities of the Study

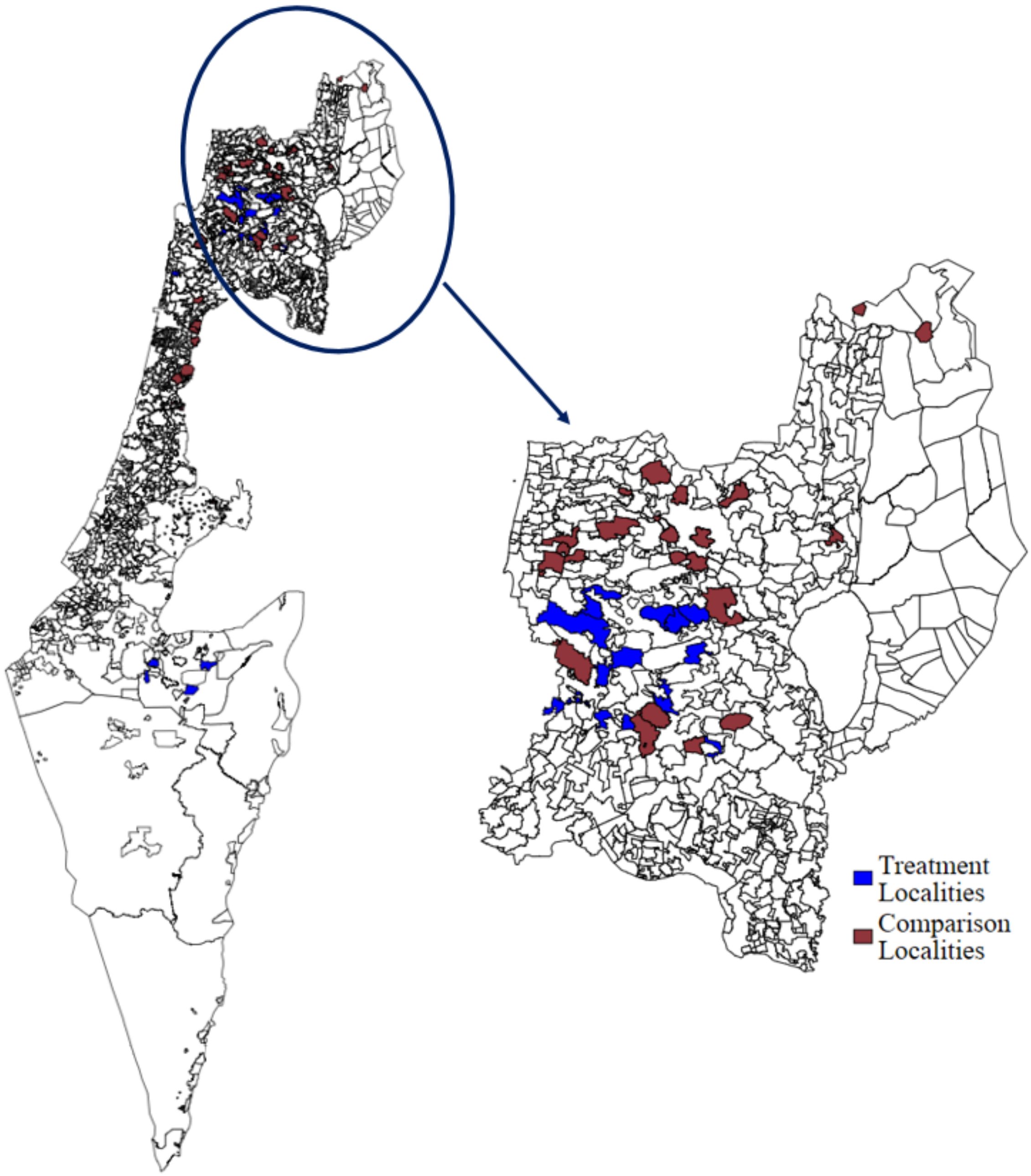
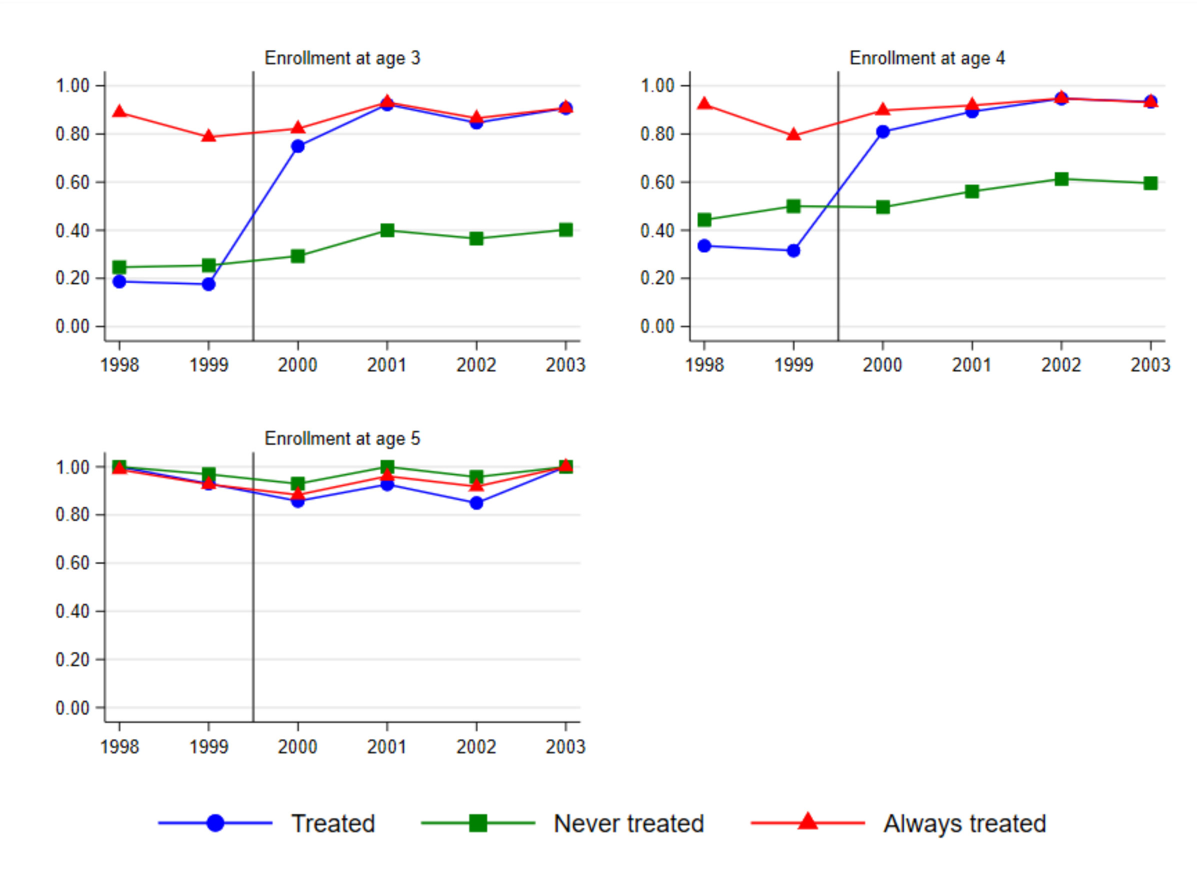


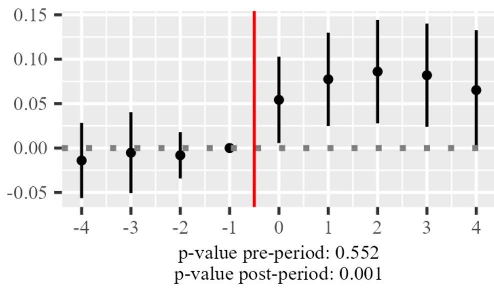
Figure 2: Preschool Enrollment in the Localities of the Study (Northern District) – 1998-2003



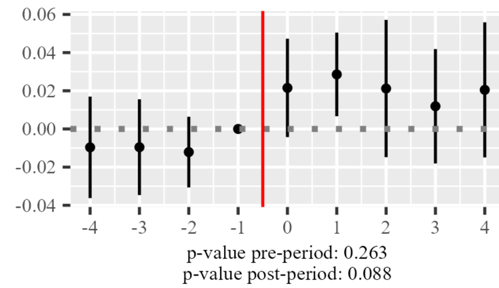
Notes: The figure shows preschool enrollment rates of Arab children by year in different groups of localities in the Northern district, according to their treatment status. The analysis is based on aggregated enrollment data and population count data by locality and year provided by the Central Bureau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

Figure 3: Event-Study Estimates of the Effects of Universal Preschool

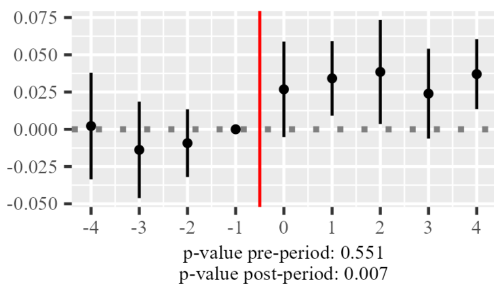
(a) High school performance z-score



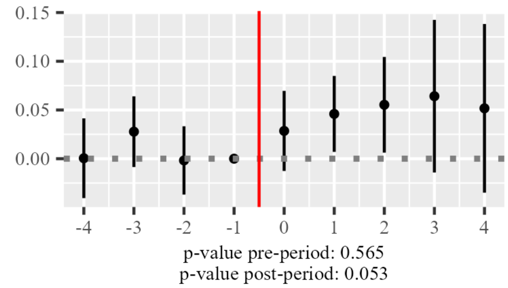
(b) Graduated from high school



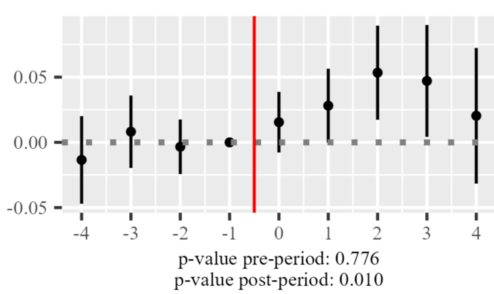
(c) Took matriculation exams



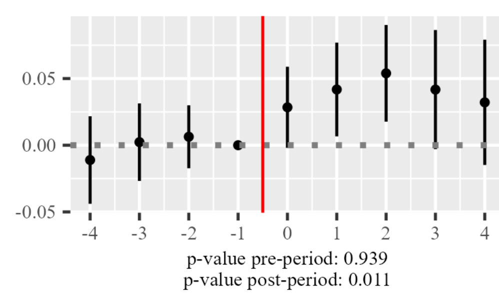
(d) Matriculation certificate



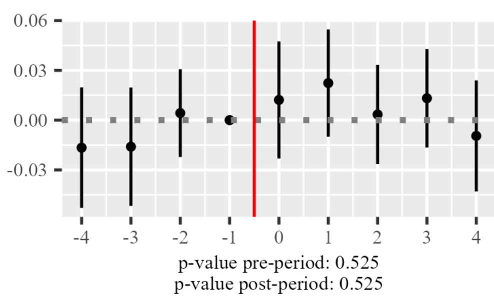
(e) University-eligible certificate



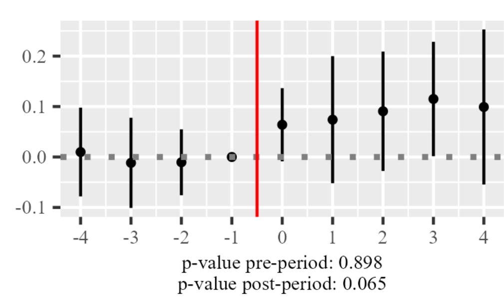
(f) 4+ English units



(g) 4+ math units



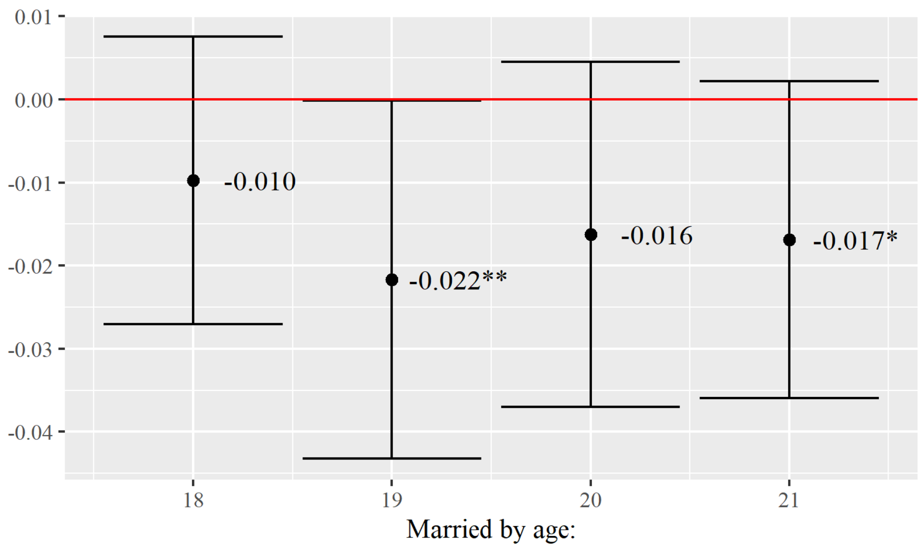
(h) Number of science subjects



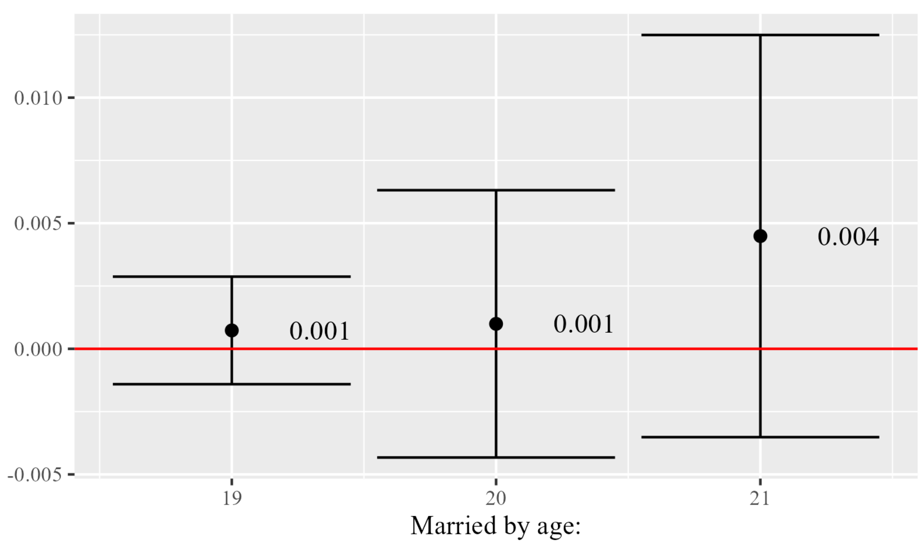
Notes: The figures plot the pretreatment and postreatment effects along 95 percent confidence intervals on high school outcomes, based on an event-study specification. The x-axis represents the years before and after the Law implementation. Year zero represents the first year of the Law implementation. 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 sample includes Israeli Arabs from localities in the north, born between 1991-1999. Standard errors are clustered at the locality level. p-values for a joint test of significance for the coefficients in the pre or the post-reform period are reported at the bottom of each subfigure.

Figure 4: Impact of Universal Preschool on Individuals' Probability of Marrying at Young Age

(a) Women



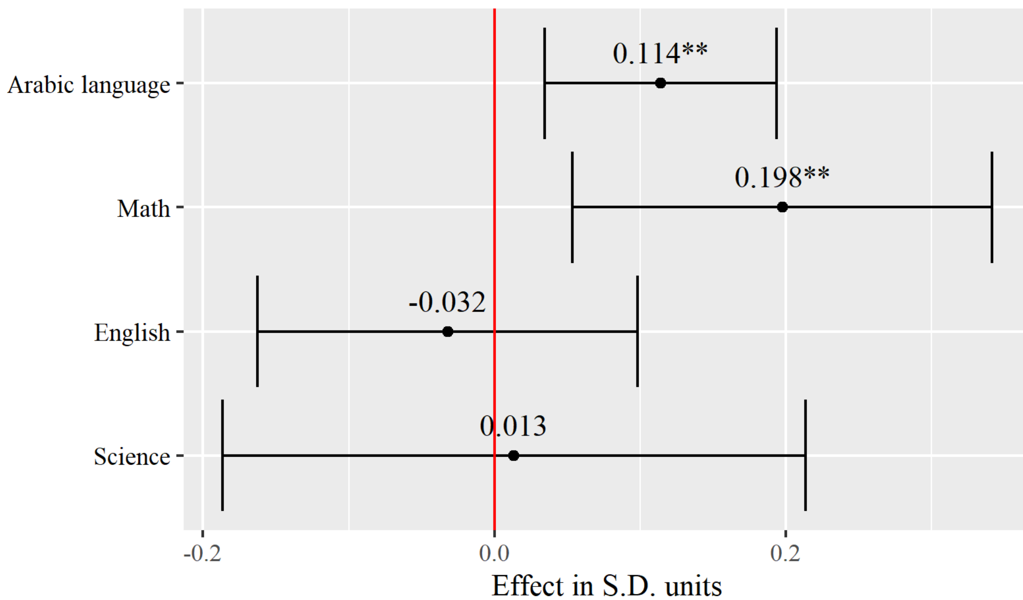
(b) Men



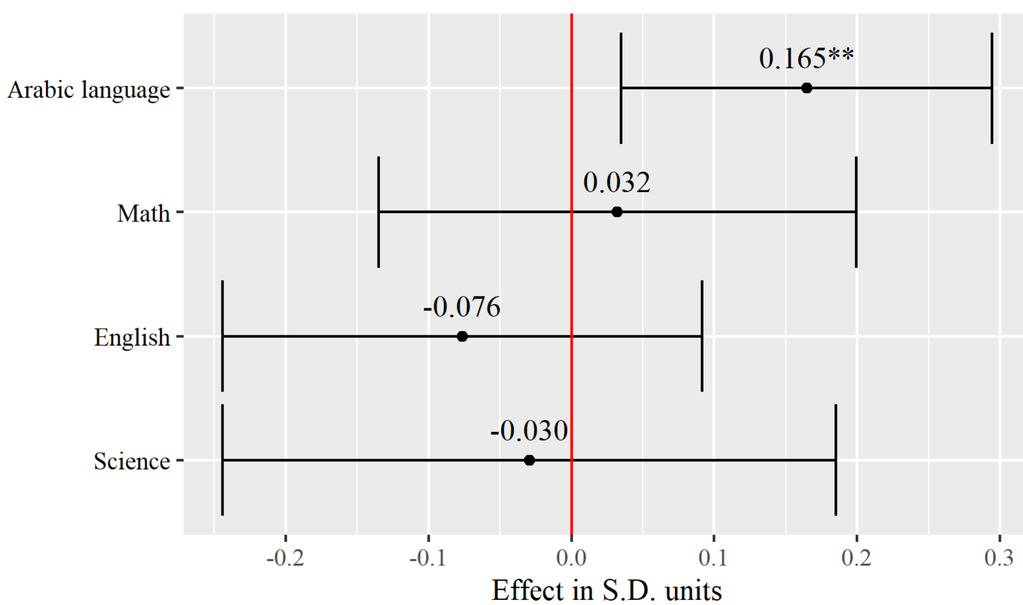
Notes: The figure reports DID estimates and 95 percent confidence intervals of the effects of universal preschool on the probability of marrying by age 18 (only for women), 19, 20, and 21, based on the specification in equation (1). The specification includes locality and cohort fixed effects and controls 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. *p<.0.10, **p<0.05, *** p<0.01.

Figure 5: Impact of Universal Preschool on 5th and 8th Grade Test Scores

(a) 5th Grade



(b) 8th Grade



Notes: The figure DID estimates and 95 percent confidence intervals of the effects of universal preschool on test scores in 5th and 8th grade. Test scores were transformed into z-scores. The specification includes locality and cohort-by-test-year fixed-effect 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. p*<0.10, **p<0.05, *** p<0.01

Table 1: Impact of Universal Preschool on High School Achievement

Dependent Variable	Full Sample (1)	Boys (2)	Girls (3)
High school performance z-score	0.080***	0.074***	0.087***
	(0.020)	(0.026)	(0.024)
	-0.058	*-0.297*	*0.196*
Graduated from high school	0.028**	0.030	0.026**
	(0.012)	(0.019)	(0.011)
	0.802	*0.690*	*0.920*
Took matriculation exams	0.037***	0.050***	0.023**
	(0.011)	(0.016)	(0.010)
	0.763	*0.635*	*0.899*
Matriculation certificate	0.043*	0.022	0.067**
	(0.023)	(0.022)	(0.030)
	0.396	*0.278*	*0.522*
University-eligible certificate	0.035***	0.024*	0.048***
	(0.013)	(0.013)	(0.018)
	0.287	*0.190*	*0.390*
4+ English units	0.040**	0.029	0.054***
	(0.016)	(0.017)	(0.020)
	0.364	*0.252*	*0.482*
4+ math units	0.015*	0.018**	0.013
	(0.009)	(0.007)	(0.014)
	0.197	*0.140*	*0.258*
Number of science subjects	0.092**	0.098**	0.089*
	(0.041)	(0.038)	(0.046)
	0.688	*0.484*	*0.904*
Number of observations	84,425	43,345	41,080

Notes: This table shows DID estimates of the impact of universal preschool on various educational outcomes. 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. The high school performance z-score (first row), is an average of all standardized individual outcomes. *p<0.10, **p<0.05, *** p<0.01

Table 2: Impact of Universal Preschool on Psychometric Test Performance

Dependent Variable	Full Sample (1)	Boys (2)	Girls (3)	Dependent Variable	Full Sample (1)	Boys (2)	Girls (3)
Took the psychometric exam	0.028***	0.037***	0.020*	Took the psychometric exam by age 19	0.033***	0.044***	0.024**
	(0.008)	(0.009)	(0.010)		(0.008)	(0.009)	(0.010)
	0.389	*0.252*	*0.534*		0.350	0.214	0.494
Total score				Quantitative score			
above first quartile (≥400)	0.022***	0.033***	0.010	above first quartile (≥85)	0.025***	0.034***	0.017**
	(0.006)	(0.007)	(0.009)		(0.005)	(0.006)	(0.008)
	0.269	*0.181*	*0.362*		0.284	0.197	0.377
above second quartile (≥470)	0.017***	0.021***	0.013	above second quartile (≥99)	0.020***	0.024***	0.016**
	(0.006)	(0.006)	(0.009)		(0.005)	(0.006)	(0.007)
	0.177	*0.126*	*0.230*		0.188	0.142	0.238
above third quartile (≥580)	0.009	0.015***	0.002	above third quartile (≥119)	0.011**	0.021***	0.001
	(0.005)	(0.005)	(0.008)		(0.005)	(0.005)	(0.007)
	0.069	*0.051*	*0.088*		0.088	0.071	0.106
Verbal score				English score			
above first quartile (≥80)	0.016**	0.030***	0.002	above first quartile (≥78)	0.025***	0.032***	0.018
	(0.006)	(0.007)	(0.009)		(0.008)	(0.008)	(0.011)
	0.269	*0.171*	*0.373*		0.219	0.147	0.295
above second quartile (≥93)	0.017**	0.025***	0.009	above second quartile (≥88)	0.021***	0.026***	0.016
	(0.006)	(0.007)	(0.009)		(0.008)	(0.008)	(0.011)
	0.188	*0.122*	0.258		0.149	0.103	0.197
above third quartile (≥109)	0.011**	0.014**	0.009	above third quartile (≥107)	0.005	0.008	0.001
	(0.005)	(0.006)	(0.007)		(0.007)	(0.005)	(0.011)
	0.094	*0.064*	0.125		0.071	0.050	0.092
Number of observations	84,425	43,345	41,080	Number of localities	37	37	37

Notes: This table shows DID estimates of the impact of universal preschool on participation and achievement in the Israel psychometric exam. 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. * p<0.10, **p<0.05, ***p<0.01

Table 3: Impact of Universal Preschool on Postsecondary Education

	Ever Enrolled			Enrolled by Age 19		
	Full Sample	Boys	Girls	Full Sample	Boys	Girls
Dependent Variable	(1)	(2)	(3)	(4)	(5)	(6)
Postsecondary enrollment	0.053***	0.065***	0.041***	0.034***	0.024***	0.044***
	(0.010)	(0.014)	(0.014)	(0.006)	(0.006)	(0.011)
	0.332	*0.245*	*0.423*	*0.157*	*0.103*	*0.214*
Enrolled at academic institution	0.040***	0.044***	0.036**	0.028***	0.015***	0.041***
	(0.008)	(0.009)	(0.013)	(0.006)	(0.005)	(0.011)
	0.262	*0.148*	*0.384*	*0.121*	*0.057*	*0.189*
University (first tier)	0.040***	0.033***	0.048***	0.029***	0.017***	0.041***
	(0.006)	(0.007)	(0.009)	(0.004)	(0.004)	(0.007)
	0.149	*0.088*	*0.212*	*0.068*	*0.036*	*0.102*
Second tier academic institution	0.023***	0.022***	0.025***	0.005	-0.001	0.011
	(0.005)	(0.004)	(0.008)	(0.004)	(0.003)	(0.007)
	0.071	*0.057*	*0.086*	*0.024*	*0.017*	*0.031*
Teacher training institution	-0.014**	-0.005**	-0.024**	-0.006*	-0.001	-0.011*
	(0.006)	(0.002)	(0.011)	(0.003)	(0.001)	(0.006)
	0.067	*0.015*	*0.122*	*0.030*	*0.004*	*0.057*
Enrolled at vocational postsecondary institution	0.020***	0.030***	0.010**	0.007**	0.009**	0.004
	(0.007)	(0.010)	(0.005)	(0.003)	(0.004)	(0.003)
	0.080	*0.109*	*0.051*	*0.036*	*0.046*	*0.026*
Number of localities	37	37	37	37	37	37
Number of observations	74,424	38,184	36,240	74,424	38,184	36,240

Notes: This table shows DID estimates of the impact of universal preschool on postsecondary enrollment. 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 treated localities appear in italics. Standard errors (in parentheses) are clustered at the locality level. *p<0.10, **p<0.05, *** p<0.01

Table 4: Impact of Universal Preschool on Juvenile Crime

Dependent Variable	Full Sample (1)	Boys (2)	Girls (3)
Any juvenile criminal offense	-0.015**	-0.030***	-0.000
	(0.006)	(0.011)	(0.001)
	0.087	*0.165*	*0.004*
Security/order criminal offense	-0.004	-0.008	-0.000
	(0.004)	(0.007)	(0.001)
	0.046	*0.088*	*0.002*
Life/body criminal offense	-0.011***	-0.022***	0.001
	(0.003)	(0.006)	(0.001)
	0.047	*0.090*	*0.002*
Sex/property criminal offense	-0.008*	-0.017**	-0.000
	(0.004)	(0.008)	(0.001)
	0.040	*0.077*	*0.001*
Other criminal offense	-0.002	-0.004	-0.000
	(0.003)	(0.006)	(0.000)
	0.016	*0.030*	*0.001*
Number of localities	37	37	37
Number of observations	84,425	43,345	41,080

Notes: This table shows DID estimates of the impact of universal preschool on the probability of having a juvenile criminal record. 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. *p<0.10, **p<0.05, ***p<0.01

Table 5: Impact of Universal Preschool - Family Fixed Effects Model

Dependent Variable	Locality FE Main Sample (1)	Locality FE Siblings Sample (2)	Family FE Siblings Sample (3)
High school performance z-score	0.080***	0.079***	0.082***
	(0.020)	(0.019)	(0.027)
	-0.058	*-0.046*	*-0.046*
Took the psychometric exam	0.028***	0.031***	0.042***
	(0.008)	(0.007)	(0.012)
	0.389	*0.395*	*0.395*
Postsecondary enrollment by age 19	0.034***	0.035***	0.029***
	(0.006)	(0.007)	(0.010)
	0.157	*0.157*	*0.157*
Any juvenile criminal offense (men)	-0.030***	-0.038***	-0.035**
	(0.011)	(0.013)	(0.015)
	0.165	*0.173*	*0.173*
Married by age 21 (women)	-0.017*	-0.021	-0.016
	(0.009)	(0.014)	(0.025)
	0.318	*0.342*	*0.342*
Number of localities	37	37	37
Number of observations	84,425	69,556	69,556

Notes: This table shows estimates of the impact of universal preschool, comparing locality fixed effects (Columns (1) and (2)) to family fixed effects (Column (3)). All specifications include 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. *p<0.10, **p<0.05, ***p<0.01

Table 6: Heterogeneous Effects of Universal Preschool

	Mother's education		Father's education		Father's annual income		Mother's employment	
	<12	≥12	<12	≥12	< median	≥ median	Not emp.	Employed
Dependent Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
High school performance z-score	0.089***	0.050**	0.086***	0.067**	0.072***	0.083***	0.090***	0.041*
	(0.024)	(0.020)	(0.023)	(0.025)	(0.022)	(0.021)	(0.022)	(0.022)
	-0.222	*0.387*	*-0.211*	*0.311*	*-0.141*	*0.066*	*-0.111*	*0.238*
Took the psychometric exam	0.033***	0.016	0.024***	0.034**	0.019***	0.036***	0.031***	0.017
	(0.008)	(0.015)	(0.007)	(0.013)	(0.007)	(0.013)	(0.007)	(0.017)
	0.306	*0.616*	*0.310*	*0.578*	*0.354*	*0.442*	*0.361*	*0.544*
Postsecondary enrollment by age 19	0.024***	0.039***	0.021***	0.056***	0.024***	0.045***	0.033***	0.039***
	(0.006)	(0.011)	(0.006)	(0.013)	(0.007)	(0.007)	(0.007)	(0.012)
	0.108	*0.292*	*0.115*	*0.258*	*0.138*	*0.186*	*0.142*	*0.240*
Any juvenile criminal offense (men)	-0.031**	-0.025**	-0.027**	-0.033***	-0.029**	-0.031***	-0.027**	-0.047***
	(0.013)	(0.009)	(0.012)	(0.010)	(0.013)	(0.010)	(0.012)	(0.015)
	0.184	*0.115*	*0.186*	*0.117*	*0.181*	*0.143*	*0.167*	*0.158*
Married by age 21 (women)	-0.010	-0.018	-0.009	-0.026	-0.034***	-0.003	-0.015	-0.021
	(0.010)	(0.012)	(0.009)	(0.020)	(0.010)	(0.012)	(0.009)	(0.023)
	0.369	*0.179*	*0.353*	*0.235*	*0.342*	*0.283*	*0.334*	*0.229*
Number of localities	37	37	37	37	37	37	37	37
Number of observations	50,642	33,643	51,442	32,559	42,212	42,213	65,667	18,758

Notes: This table shows DID estimates of the impact of universal preschool on various 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. *p<0.10, **p<0.05, ***p<0.01

Table 7: Heterogenous Effects of Universal Preschool by Predicted Outcomes

	Level of Predicted Outcome		
	Low	Medium	High
Dependent Variable	(1)	(2)	(3)
High school performance z-score	0.080**	0.104***	0.034*
	(0.030)	(0.028)	(0.019)
	-0.457	*0.020*	*0.574*
Took the psychometric exam	0.026***	0.029**	0.015
	(0.009)	(0.013)	(0.014)
	0.172	*0.421*	*0.727*
Postsecondary enrollment by age 19	0.018***	0.027***	0.049***
	(0.006)	(0.009)	(0.014)
	0.064	*0.151*	*0.344*
Any juvenile criminal offense (men)	-0.018**	-0.021	-0.010
	(0.008)	(0.013)	(0.015)
	0.081	*0.152*	*0.204*
Married by age 21 (women)	-0.019	-0.006	-0.006
	(0.022)	(0.016)	(0.012)
	0.127	*0.289*	*0.396*
Number of localities	37		
Number of observations	84,425		

Notes: This table shows the estimated effects of universal preschool by tertiles of predicted outcomes, as defined by the pre-treatment relationship between outcomes 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. *p<0.10, **p<0.05, ***p<0.01

Table 8: Impact of Universal Preschool on Learning Environment

Dependent Variable		Dependent Variable	
Learning environment z-score	0.101***		
	(0.028)		
	-0.070		
Satisfaction with school and classroom		**Placebo outcomes: computer use**	
I enjoy school	0.042***	Computer at home	0.001
	(0.012)		(0.008)
	0.760		*0.773*
Students in my classroom help each other	0.021**	Use of computer in Arabic lessons	0.001
	(0.009)		(0.014)
	0.770		*0.370*
There are frequent disturbances in the classroom	-0.042**	Use of computer in English lessons	-0.010
	(0.020)		(0.021)
	0.752		*0.366*
Safety and security			
Teachers prevent violence/maintain discipline	0.030*	Use of computer in math lessons	-0.000
	(0.013)		(0.022)
	0.839		*0.397*
Sometimes I'm afraid to go to school because there are violent students	-0.053***	Use of computer in science lessons	-0.021
	(0.016)		(0.045)
	0.300		*0.478*
I have someone in school to consult with	0.031**		
	(0.013)		
	0.762		
Relationship with teachers			
There are good relationships between teachers and students	0.025***		
	(0.009)		
	0.791		
Sometimes teachers insult children	-0.072***		
	(0.023)	Number of localities	37
	0.464	Number of observations	144,144

Notes: This table shows DID estimates of the impact of universal preschool on various learning environment outcomes, as reflected in students' answers to the GEMS questionnaires in grades 5-9. The outcome is a dummy variable that takes the value of 1 if respondents partially to strongly agree, and 0 if they partially to strongly disagree. The learning environment z-score is a standarized average of all learning environment outcomes. The specification includes locality, cohort, year, and grade fixed effects controlling for the type of school (Arab/Druze/Bedouin). 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. *p<0.10, **p<0.05, ***p<0.01.

Table 9: Comparison to Similar Studies - Local Average Treatment Effects

Study	Type of preschool, country	Age at intervention	High school graduation: Effect	High school graduation: Baseline mean	College enrollment: Effect	College enrollment: Baseline mean	Counterfactual mode of care	Maternal education
A. Large-scale Programs								
Gray-Lobe et al. (2023)	Universal, US (Boston)	4	0.060	0.64	0.054	0.65	68% in formal care, 32% in home/informal care	n.a
Havnes and Mogstad (2011)	Universal, Norway	3-6	0.058	0.74	0.069	0.37	informal care	avg.=9.7
Deming (2009)	Head Start, US	3-5	0.086	unknown	0.057	unknown	n.a	less than high school=46%
Bailey et al. (2021)	Head Start, US	3-5	0.024	0.92	0.054	0.64	home/ informal care	less than high school=66%
This study	**Universal, Israeli Arabs**	**3-4**	**0.047**	**0.80**	**0.067**	**0.26**	**most at home**	**avg.= 9.4; less than high school=60%**
B. Targeted Programs								
Belfield et al. (2006)	Perry Preschool, US	3-5	0.168	0.60 (at age 40)			home/informal care	9.2
Campbell et al. (2012)	Abecedarian, US	0-6	0.068	0.82	0.17	0.06	25% home care, 75% alternative care arrangement	10.5
Heckman et al. (2010)	Perry Preschool, US	3-5	0.61 (girls) -0.03 (boys)	0.23 (girls) 0.51 (boys)			home/informal care	9.2
Anderson (2008) - high school Elango et al. (2016) - college	Abecederian, US	0-6	0.23 (girls) -0.10 (boys)	0.61 (girls) 0.74 (boys)	0.193	unknown	25% Home Care, 75% alternative care arrangement	10.5

Notes: This table presents a comparison of estimates on the effects of preschool education on high school graduation and college enrollment. **R**eferences for the estimates reported for each study appear in Table A20.

Table 10: Comparison to Other Educational Interventions Implemented in Israel at Older Ages

					Matriculation certificate		Postsecondary enrollment	
Study	Intervention	Target population	Age	Cost per student (2000)	Effect	Baseline mean	Effect	Baseline mean
Lavy and Schlosser (2005) Lavy et al. (2022)	Remedial education	Underperforming students at the margin of earning a matriculation certificate in low-achieving schools	15-18	$1,100	0.13	0.55	0.08 Increased college (second-tier) enrollment with no effect on university (first-tier) enrollment.	0.55
Angrist and Lavy (2009)	Monetary awards to students	Students in 39 low-achieving high schools (10 Arab schools)	15-18	$385	0.14 girls no effect for boys	0.24 all 0.29 girls 0.20 boys	No effect overall or on university (first-tier) enrollment; Increased second tier enrollment for girls in the top quartile by 0.16	0.43 (girls in top quartile of achievement distribution)
This study	**Universal preschool**	**Israeli Arabs in low SES localities**	**3-4**	**$1,400**	**0.07**	**0.4**	**0.09** **Increased enrollment in almost all institution types, including university**	**0.33**

Notes: This table compares the long-term impact of universal preschool to the impact of high schools interventions in Israel during the same period, based on the likelihood of earning a matriculation certificate and of enrolling in a postsecondary institution. Estimates of the impact of universal preschool are scaled up by the increase in preschool enrollment generated by implementation of the law (60%). The estimate of the effect of remedial education on the likelihood of earning a matriculation certificate is taken from Lavy et al. (2022), Table 2, columns 1-3, which is identical to the estimate reported in Lavy and Schlosser (2005), Table 8, columns 1-3. The baseline mean for this outcome is computed by subtracting the treatment effect (0.13) from the outcome mean of the treated group (0.681) reported in Lavy et al. (2022), Table 2, column 2 . The estimate of the impact of remedial education on postsecondary enrollment is taken from Lavy et al. (2022), Table 3, columns 1-3. The baseline mean for this outcome is computed by subtracting the treatment effect (0.08) from the outcome mean of the treated group (0.631) reported in column 2 of that table. Estimates of the effects of monetary awards on the probability of earning a matriculation certificate are based on Angrist and Lavy (2009), Table 2, columns 3 and 5, and are scaled up by treatment take-up (75%). Estimates of the effects of monetary awards on postsecondary enrollment are based on Angrist and Lavy (2009), Table 8, Panel C, and are scaled-up by treatment take-up (75%).

Short- and Long-Term Effects of Universal Preschool:

Evidence from the Arab Population in Israel

Elad DeMalach

Bank of Israel and Tel Aviv University

Analia Schlosser

Tel Aviv University, CEPR, CESifo, and IZA

ONLINE APPENDIX

Table of Contents

Appendix A – Details on Universal Preschools 2

Appendix B – Data 4

Appendix C - Assessing the Parallel Trends Assumption 7

Appendix D – Details of the Robustness Checks 8

Appendix E – Analysis of Selection into Identification in Family Fixed Effects Model 11

Appendix F – Impact on Maternal Employment and Earnings 12

Appendix Figures and Tables 14

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.

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.

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.

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

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.

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

[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).

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 $\overline{\mathrm{M}}$ times the maximum pre-treatment violation of the parallel trend. The breakdown point is $\overline{\mathrm{M}} \approx 1.1$, 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 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-

[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).

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 3 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

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

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]

Late-Treated Localities

As noted in the background section (Section 2), 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

---

[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.

[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.

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 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

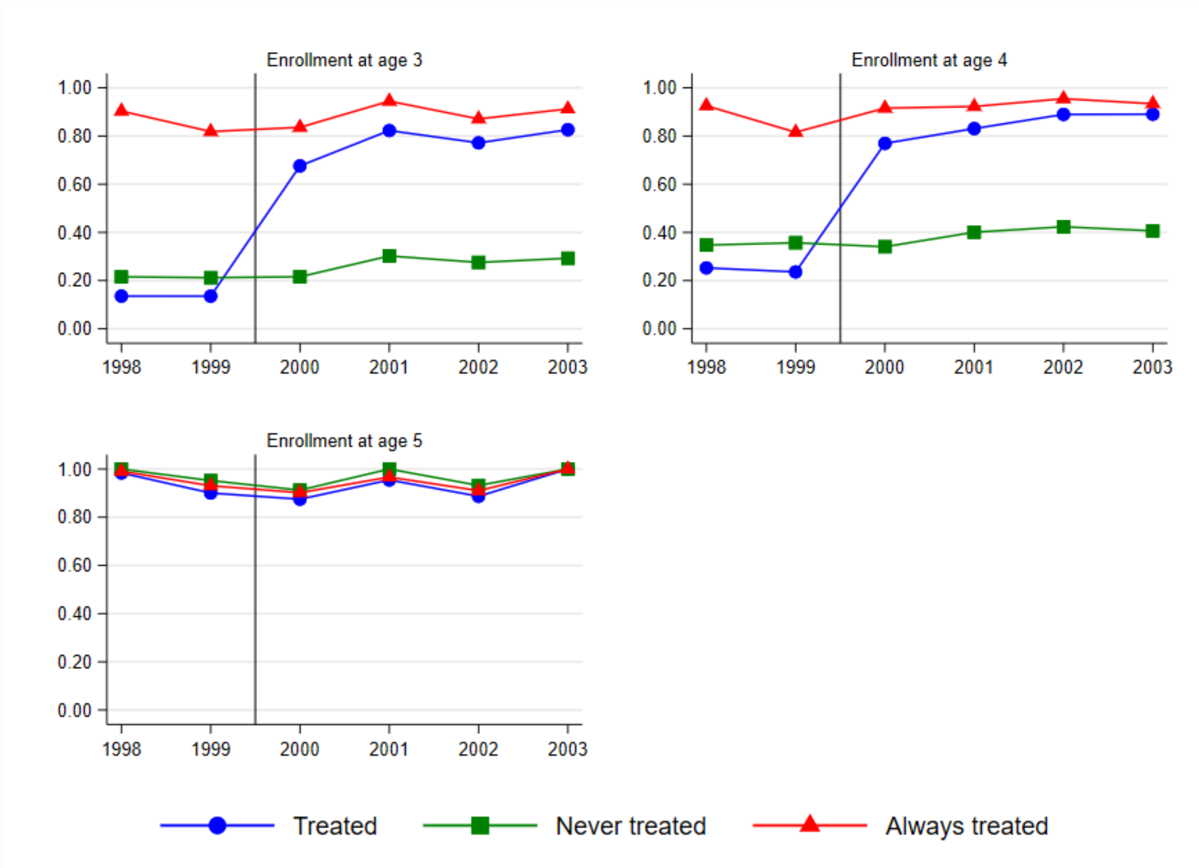
[8] We also select mothers of children aged 5 in this sample, as most children turn 5 while attending preschool (the cutoff date for entering grade 1 was around September 1).

[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).

mothers’ education (Table A19). 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.

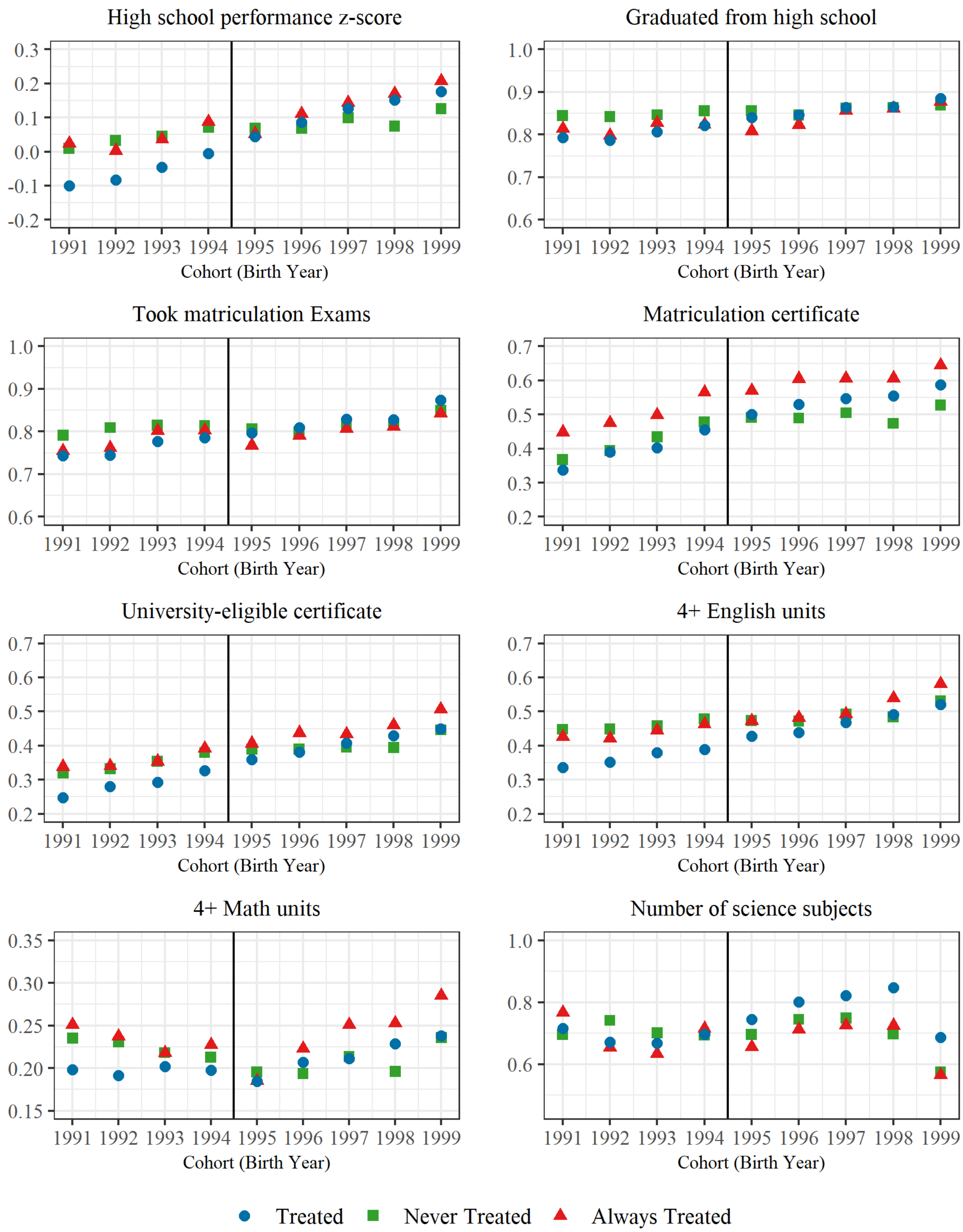
Appendix Figures and Tables

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.

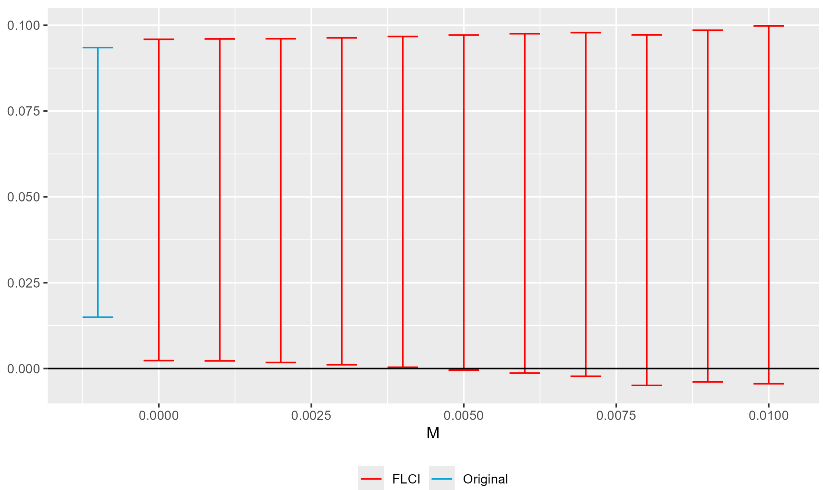
Figure A2: Unconditional cohort means, by treatment status



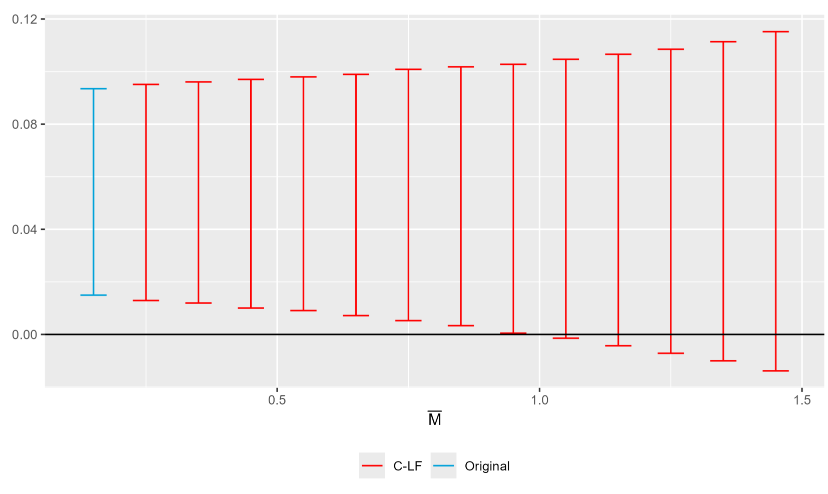
Notes: The figure shows unconditional cohort means of high school outcomes according to the locality treatment status. Treated localities received universal preschool education starting from the year 2000 (1995 cohort). Never-treated localities are those that were not included in the first phase of the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

Figure A3: Sensitivity Analysis for the Treatment Effect on High School Performance to Violations of the Parallel Trends Assumption

(a) Smoothness Restriction

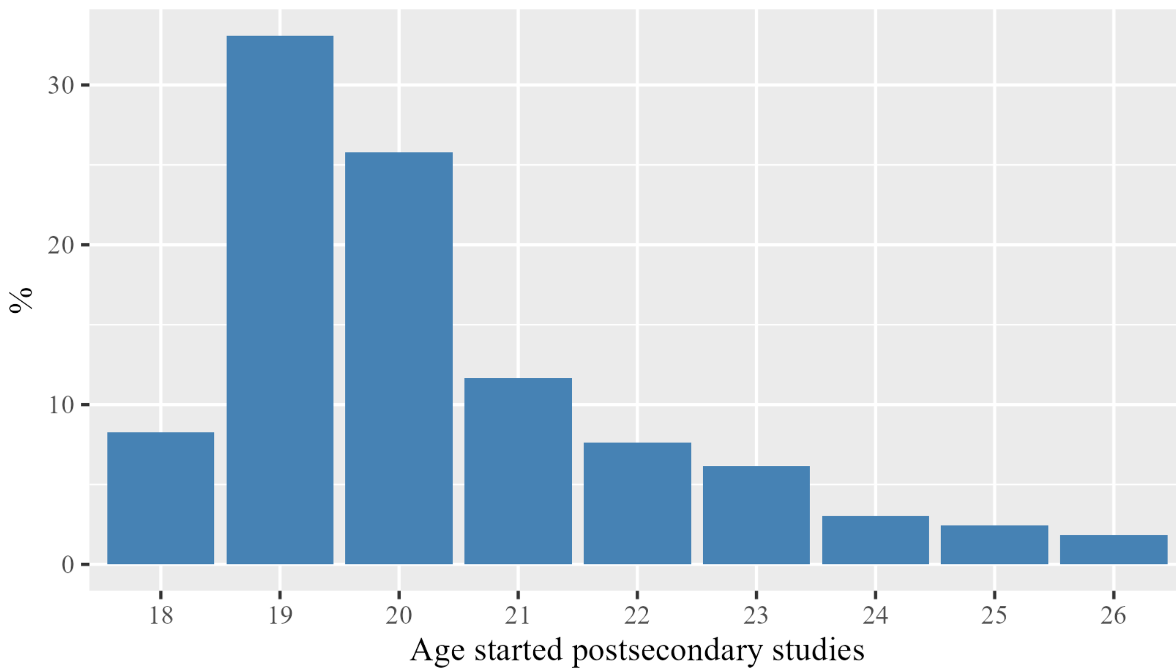


(b) Bounds on Relative Magnitude



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 $\bar{M}$ times the maximum pre-treatment violation of the parallel trend (sensitivity analysis using relative magnitudes restrictions).

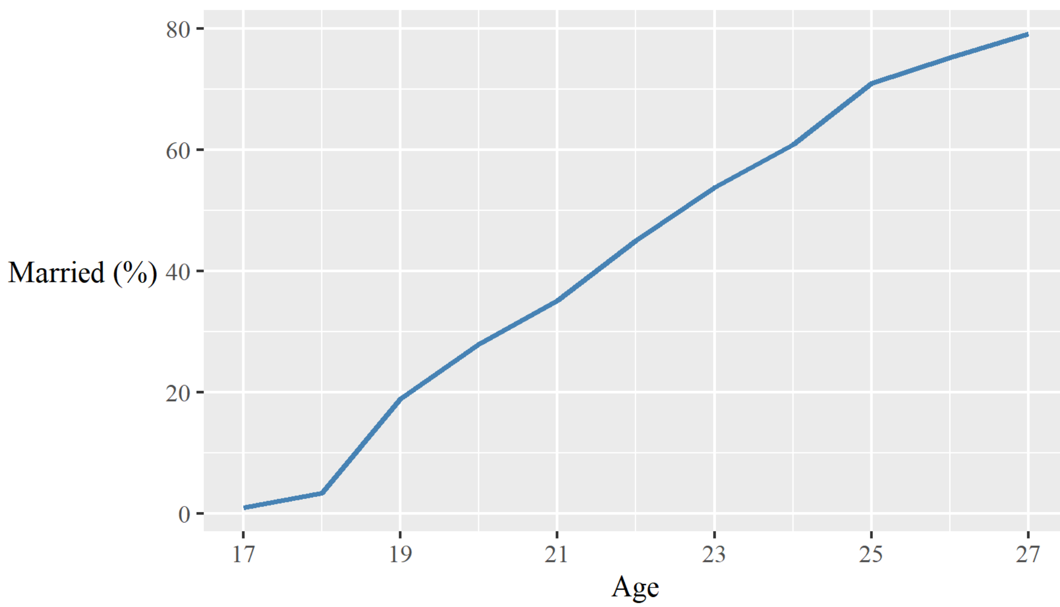
Figure A4: Age Distribution at Enrollment in Postsecondary Institutions



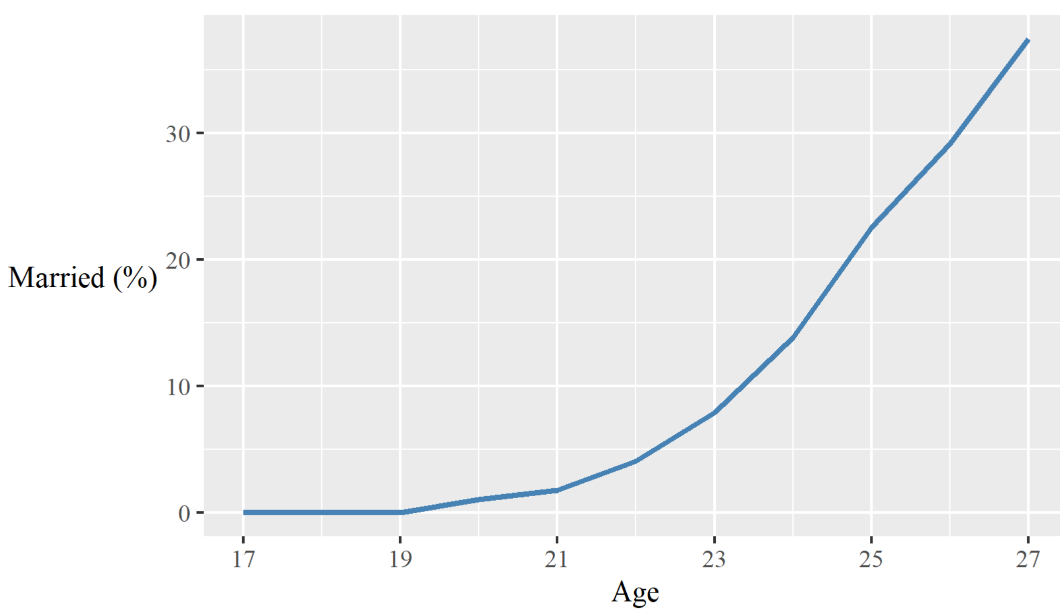
Notes: The figure reports the age distribution at first enrollment in a postsecondary education institution for the prereform cohort (born in 1991) in the localities of this study. Enrollment data is available until the 2017-2018 academic year.

Figure A5: Share of Married Individuals, by Age

(a) Women

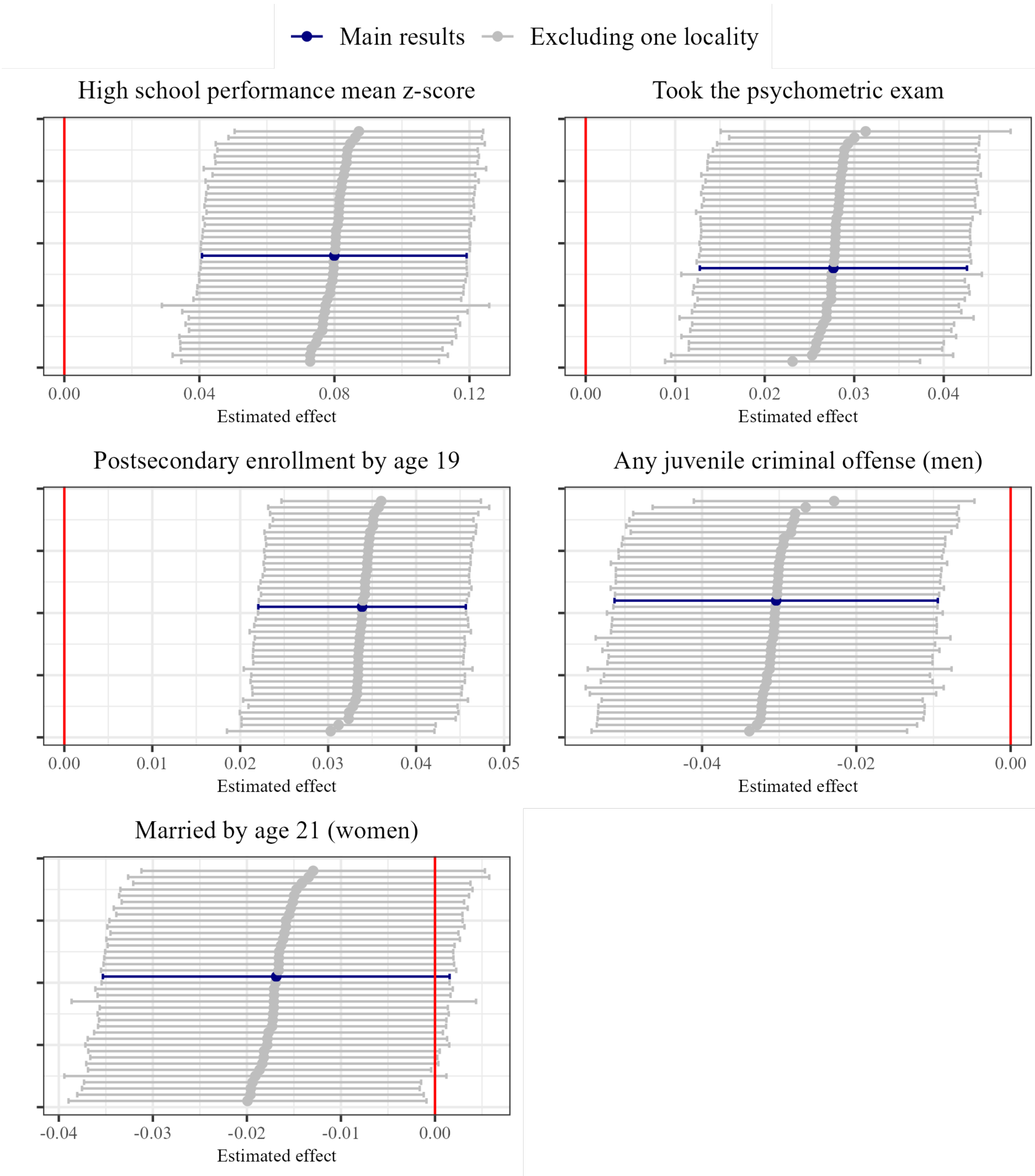


(b) Men



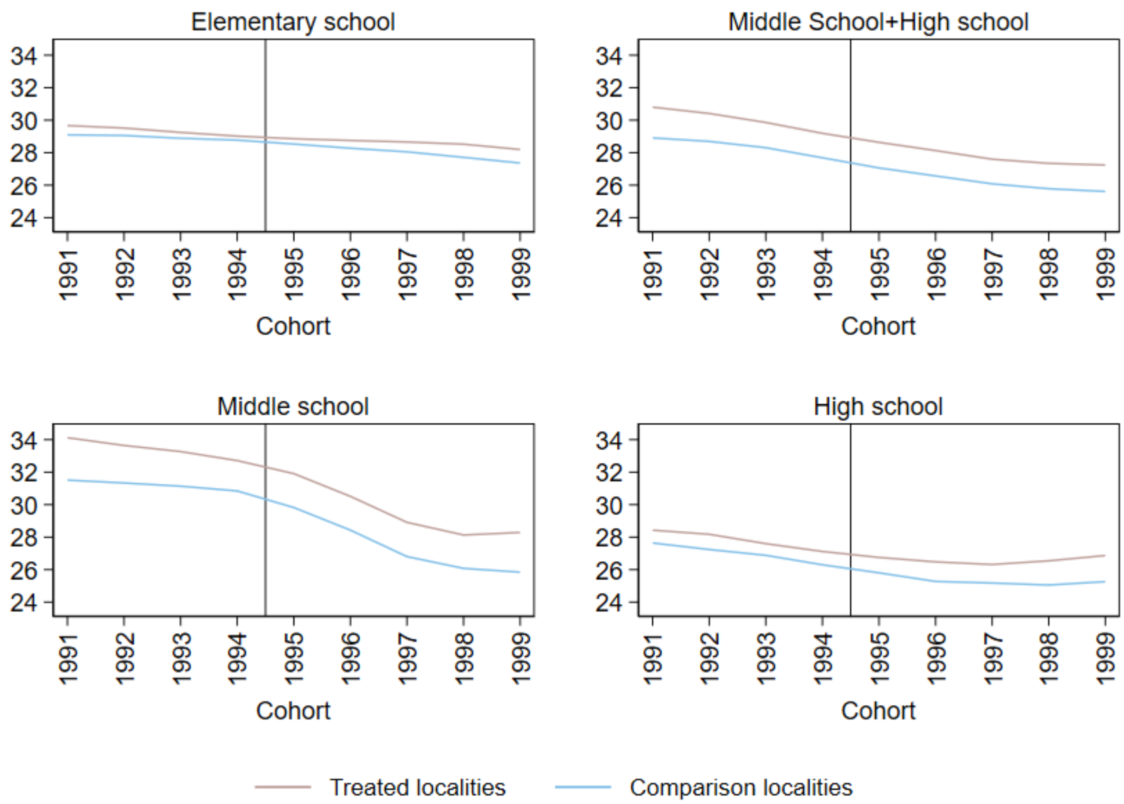
Notes: The figure plots the share of married individuals by age for the prereform cohort born (born in 1991) in the localities of this study.

Figure A6: Sensitivity Analysis of the Impact of Universal Preschool



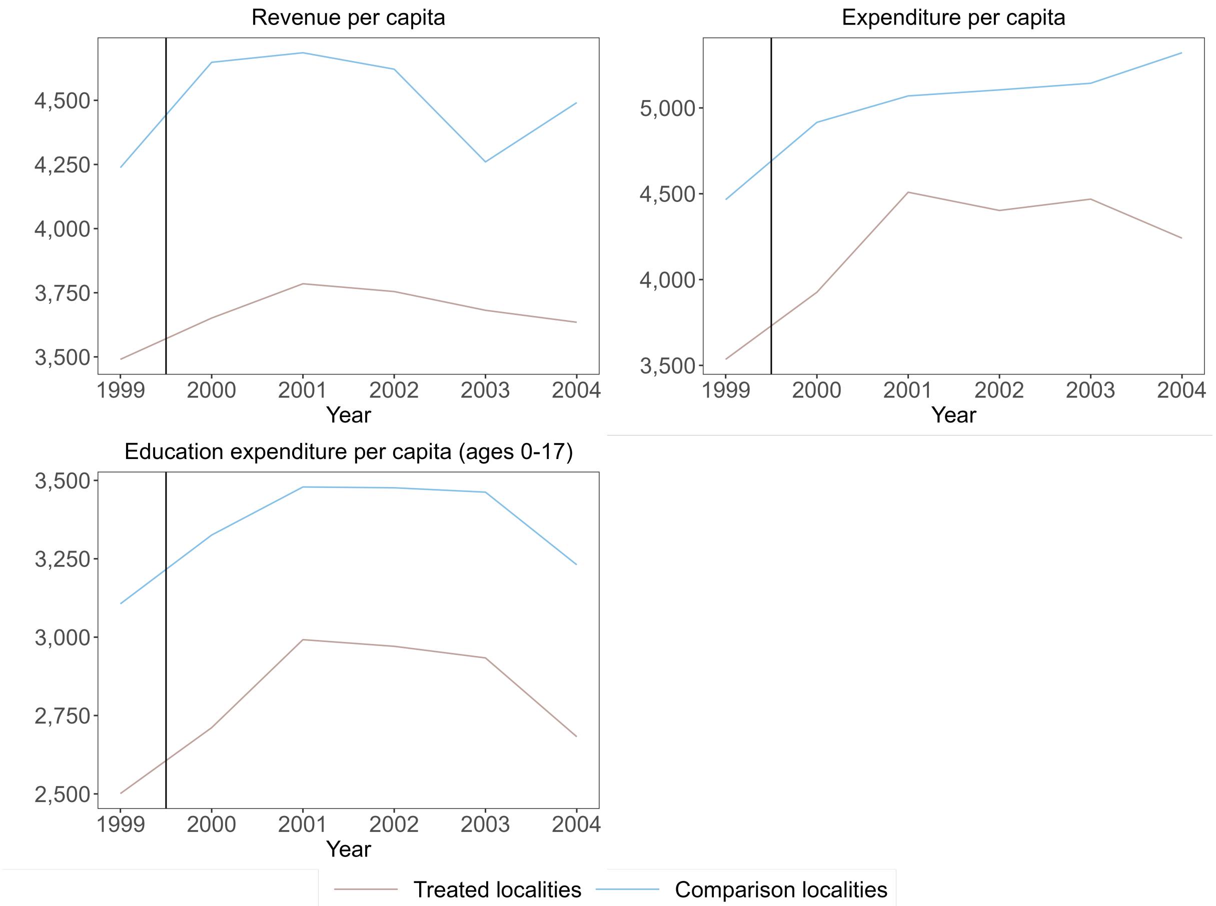
Notes: The figures plot the distribution of estimates and 95% confidence intervals of our baseline DID specification in equation (1). The blue bars represent estimates for our main sample, and the grey bars represent estimates obtained by excluding one locality from the sample at a time. The specification 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		Juvenile crime
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		
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	High school graduation, matriculation, psychometric exams, Postsecondary enrollment, Marriage	
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		
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	18.90***
	(14.71)	(12.74)	(4.74)
Families with 4 or more children (%)	0.40	0.30	0.10***
	(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). *p<0.10, **p<0.05, ***p<0.01

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. *p<0.10, **p<0.05, ***p<0.01

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	0.080***	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)	-0.030***	-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. *p<0.10, **p<0.05, ***p<0.01

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. * p<0.10, **p<0.05, *** p<0.01

Table A7: Robustness Checks - Placebo Timing of Treatment

Dependent Variable	Main results (1)	Pre-reform "placebo" effect (2)
High school performance z-score	0.080***	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)	-0.030***	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. *p<0.10, **p<0.05, ***p<0.01

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	0.080***	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***	0.020***	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***	0.030***	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)	-0.030***	-0.023**	-0.040***	-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**	-0.020*
	(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. *p<0.10, **p<0.05, ***p<0.01

Table A9: Robustness Checks - Wild Cluster Bootstrap

Dependent Variable	Main results (1)	Wild cluster bootstrap (2)
High school performance z-score	0.080***	0.080***
	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)	-0.030***	-0.030***
	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. *p<0.10, **p<0.05, ***p<0.01

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. *p<0.10, **p<0.05, ***p<0.01

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	0.080***	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	0.040**	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. *p<0.10, **p<0.05, ***p<0.01

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 (1)	Preschool enrollment at age 4 (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 (1)	Postsecondary enrollment by age 19 (2)	Took the psychometric exam (3)	Any juvenile criminal record (men) (4)	Married by age 21 (women) (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 Low/Median Predicted Outcome	0.057***	-0.027*	0.015	-0.016	-0.011
	(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. *p<0.10, **p<0.05, ***p<0.01

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**	0.060***	0.018
	(0.014)	(0.017)	(0.017)
	0.114	*0.294*	*0.629*
4+ English units	0.038**	0.070***	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. *p<0.10, **p<0.05, ***p<0.01

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. *p<0.10, **p<0.05, ***p<0.01

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. *p<0.10, **p<0.05, ***p<0.01

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	
	Mothers of Children Aged 3-5	Other Mothers	Mothers of Children aged 3-5	Other Mothers	Mothers of Children aged 3-5	Other Mothers
Dependent Variable	(1)	(2)	(3)	(4)	(5)	(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. *p<0.10, **p<0.05, ***p<0.01

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

	Age 3 (1)	Age 4 (2)	Age 5 (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. *p<0.10, **p<0.05, ***p<0.01

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.		

Note: 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.