Starting Together, Diverging Later? Gender Differences in Universal Pre-K's Long-Term Effects

Assaf Kott*

September 2, 2026

Abstract

Evaluations of early-childhood programs often find that boys and girls are affected differently, yet little is known about the origins of this heterogeneity. I study this question in the context of Israel's expansion of universal public pre-K, focusing on the Arab population and using administrative data that link each child's enrollment at age three to school progression and high school outcomes. Two features of the setting make this question tractable: The alternative to starting pre-K at age three is overwhelmingly home care; and a rapid expansion with rationed slots generated within-family variation in attendance, which I use in a family fixed-effects model. I find that pre-K at age three improves school progression but has limited effects on high school achievement. These effects differ sharply by sex: Girls benefit substantially more than boys across progression outcomes. I test two classes of explanations for this gap: differences in developmental readiness and differences in the home-care counterfactual. The evidence is inconsistent with the first—relatively younger boys, presumably the least ready, benefit more, not less—and more consistent with the second: A significant girl advantage appears in families where gender norms and son preference are more likely to operate, but not in their counterparts.

JEL Codes: I21, I26, I28, J13, J16, J18

Keywords: Children, Preschool, UPK, Gender Gap

*Department of Economics, Ben-Gurion University of the Negev, kotta@bgu.ac.il. This paper benefited from discussions with Anna Aizer, Chien-Tzu Cheng, Danny Cohen-Zada, Naomi Gershoni, Ada González-Torres, and Marcella Mello, as well as participants at the Workshop on the Economics of Education at Valle Nevado, EWMES 2025, LESE 2026, SEHO 2026, and seminars at Ben-Gurion University and Brown. I thank Israel's Ministry of Education for granting access to the data. All errors are my own.

1 Introduction

A broad consensus across multiple disciplines holds that shocks and interventions during infancy and early childhood can have long-lasting effects on human capital formation and later-life outcomes (Knudsen et al., 2006; Almond, Currie, and Duque, 2018). This recognition has generated substantial interest among both researchers and policymakers in the public provision of pre-K.[1] At the same time, evaluations of early-childhood programs in general—and of universal pre-K in particular—often find that boys and girls do not benefit equally: Some studies document larger gains for girls (Felfe, Nollenberger, and Rodríguez-Planas, 2015; Havnes and Mogstad, 2011), while others show larger gains for boys (Blanden et al., 2022; Gray-Lobe, Pathak, and Walters, 2023). Little is known, however, about the origins of this heterogeneity.

In this study, I estimate the long-term effects of attending universal public pre-K (UPK) at age three and explore the sources of the gender differences in these effects. I focus on the Arab population in Israel, where two features make this question tractable. The first is a clear counterfactual: The alternative to public pre-K in this setting is overwhelmingly home care. This is important because a gender gap in a program's effects—like any treatment-effect heterogeneity—can arise for two distinct reasons: Boys and girls may respond differently to the program itself, or they may face different environments in its absence (Kline and Walters, 2016; Kottelenberg and Lehrer, 2017). When a program crowds out a mix of home care and private arrangements, these two explanations are difficult to separate; here, pre-K replaces home care, so I can test both empirically.

The second feature that makes this question tractable is the research design: I study a period of rapid expansion in the supply of public pre-K, during which slots for three-year-olds were rationed. I show that these supply conditions generated variation in pre-K availability within households, and I use this variation in a family fixed-effects model to identify the treatment effect by comparing siblings who attended pre-K at age three with siblings who did not. This approach eliminates bias from family-level selection into the program. Indeed, family fixed-effects models are commonly used in the evaluation of preschool programs—a literature I review shortly—and have also been employed in a broad range of other settings.[2]

[1] Additional policy motivations often include increasing female labor force participation and raising fertility rates.

[2] These include the long-term effects of neighborhoods and housing programs (Chetty and Hendren, 2018; Pollakowski et al., 2022), intergenerational consequences of teenage motherhood (Aizer, Devereux, and Salvanes, 2022), school peer effects (Bertoni, Brunello, and Cappellari, 2020; Figlio et al., 2024; Gazze, Persico, and Spirovska, 2024),

A natural concern with sibling comparisons is that unobserved factors may systematically influence parents' decisions to enroll some of their children in pre-K but not others: As noted by Deming (2009), "*Something* is driving differences in participation among siblings." In my setting, however, parental discretion was limited: Because slots were rationed, whether a given child could attend at age three depended largely on the timing of the local rollout relative to the child's birth cohort. Consistent with this, approximately 80% of within-family switches in pre-K attendance coincide with the expansion, and the remaining switches disproportionately involve younger siblings born late in the year, as would be expected under slot rationing—and I control flexibly for month of birth. My results are robust to excluding families whose switches do not coincide with the expansion, as well as to alternative municipality samples and model specifications. In addition, I show that within-family variation in enrollment is uncorrelated with predetermined child characteristics (such as sex, birth order, and birth weight) and time-varying household conditions (namely, maternal employment at age three). Finally, I find no evidence that one sibling's pre-K enrollment affects another sibling's outcomes, which alleviates concerns about violations of the Stable Unit Treatment Value Assumption.

I find that starting pre-K at age three rather than at age four significantly improves markers of school progression: Children who attended pre-K are 1.9 percentage points more likely to start first grade on time and 2.4 percentage points more likely to graduate from high school on time. In contrast, I find limited evidence that pre-K improves educational achievement: The effects on obtaining a high school diploma (Bagrut) and on performance in key subjects such as mathematics and English are small and statistically nonsignificant. These pooled effects, however, differ sharply by sex: Girls benefit substantially more than boys across the progression outcomes (joint test for equal effects: $p < 0.001$)—for on-time high school graduation, the effect for girls is 4.2 percentage points, compared with a statistically insignificant 0.6 percentage points for boys. Notably, pre-K attendance significantly increases girls' likelihood of ever completing high school by 3.0 percentage points—a finding not observed in the pooled sample.

In the second part of the analysis, I test two explanations for this gender gap. The first is developmental readiness: Because boys' skills develop at a slower rate than girls'—especially

and various health conditions and interventions (Karbownik and Wray, 2025; Lleras-Muney, Price, and Yue, 2022; Miller, Wherry, and Aldana, 2022).

social-emotional skills (Else-Quest et al., 2006; Montroy et al., 2016)—boys may be less prepared to benefit from the pre-K environment (Crockenberg, 2003), with relatively younger boys presumed to be the least ready of all. However, contrary to this hypothesis, I find that relatively younger boys (those born later in a given year) benefit significantly more than relatively older boys.[3] The second explanation builds on the counterfactual logic above: In a population where traditional gender norms are prevalent (Yassine-Hamdan and Strate, 2020), the home environment that pre-K replaces may differ for boys and girls. In line with this hypothesis, the girl advantage appears where gender norms predict the largest differences in home care: I find a significant girl advantage in families with low maternal education but not in families with high maternal education, and a significant girl advantage in families raising both boys and girls but not in families with same-sex siblings. Taken together, the evidence is inconsistent with a readiness-based explanation and more consistent with the home-counterfactual hypothesis.

This paper makes two key contributions to the literature. First, it sheds light on the origins of gender differences in the returns to early education. Although such differences are documented across many settings, their sources remain unclear, partly because most existing estimates come from exposure designs, in which a gender gap can reflect differences in take-up as well as differences in treatment effects; observing attendance allows me to separate the two.[4] The two explanations I test correspond to two classes the literature proposes: differences in boys' and girls' human-capital production functions—boys develop social-emotional skills more slowly (Bertrand and Pan, 2013; Crockenberg, 2003; Deming and Dynarski, 2008; Reeves, 2022), and preschool environments reward precisely those skills (Magnuson et al., 2016; Fidjeland et al., 2023)—and differences in the counterfactual care environment (Kline and Walters, 2016; Kottelenberg and Lehrer, 2017), which gender norms and son preference shape through parental time and resource allocation (Dahl and Moretti, 2008; Blau et al., 2020; Baker and Milligan, 2016; Barcellos, Carvalho, and Lleras-Muney, 2014; Jayachandran and Pande, 2017; Dossi et al., 2021) and through children's daily experiences at home (Boxberger and Reimers, 2019; Tandon, Zhou, and Christakis, 2012).[5] To my knowledge,

[3] The eligibility cutoff date for age-three pre-K (and other school-entry rules) was set according to the Hebrew calendar, typically falling in December (Attar and Cohen-Zada, 2018).

[4] Havnes and Mogstad (2015) write: "A limitation of the reduced form parameters is that we cannot tell whether differential effects of the child care reform operate through differences in child care take-up or potentially heterogeneous impacts of uptake (depending on the quality of the child care center and the counterfactual form of care)" (p. 106).

[5] These norms are particularly salient in my setting, where girls are less likely to play outdoors and more likely to participate in household chores (UNDP, 2006; El Feki, Heilman, and Barker, 2017).

this is the first study to test these competing explanations within a single research design.

My second contribution is to the evidence on the long-term effects of UPK attendance. While numerous studies document the short- and medium-term effects of public pre-K (Gormley and Gayer, 2005; Berlinski, Galiani, and Manacorda, 2008; Fitzpatrick, 2008; Berlinski, Galiani, and Gertler, 2009; Kottelenberg and Lehrer, 2017; Blanden et al., 2016; Lipsey, Farran, and Durkin, 2018; Cornelissen et al., 2018; Weiland et al., 2020; Cascio, 2023; Humphries et al., 2024), examining long-term treatment effects has proven much harder.[6] Because data that jointly observe pre-K attendance and long-run outcomes are rare, existing research has largely estimated the long-run effects of exposure to pre-K programs rather than the effects of attendance (Havnes and Mogstad, 2011; Felfe, Nollenberger, and Rodríguez-Planas, 2015; Baker, Gruber, and Milligan, 2019; DeMalach and Schlosser, 2026; Berlinski et al., 2025), and evidence on the long-run treatment effects of UPK participation remains comparatively limited (Gray-Lobe, Pathak, and Walters, 2023; Akee and Clark, 2024).[7] This paper also builds on the strand of the literature that uses family fixed effects to identify the treatment effects of public preschool programs (Currie and Thomas, 1995; Berlinski, Galiani, and Manacorda, 2008; Deming, 2009; Pages et al., 2020; Miller, Shenhav, and Grosz, 2023).[8] To my knowledge, no prior study has implemented this approach in a UPK context using administrative data. These two features—a universal program and the administrative data—yield a large sample of siblings, which permits the granular analysis of treatment-effect heterogeneity at the heart of the paper; the administrative data also do not rely on retrospective reporting of pre-K attendance, as many surveys do, which could be prone to misclassification (Cascio, 2021).

This paper is closely related to DeMalach and Schlosser (2026, henceforth DS), who study pre-K in the same setting—Arab municipalities in Israel—and, using a difference-in-differences design, estimate the intent-to-treat effects of local access to pre-K at ages three and four. The two designs answer complementary questions: The DS estimates capture the full effect of the expansion on

[6] Examining long-term outcomes is essential for two reasons: First, research on early-childhood programs frequently shows a fade-out of effects in the medium term followed by a re-emergence in adulthood (Currie and Thomas, 1995; Heckman, Pinto, and Savelyev, 2013; Bruhn and Emick, 2023). Second, standardized test scores alone might not capture the full spectrum of skills that matter for adult human capital (Jackson, 2018).

[7] The literature on Head Start's long-term consequences is more advanced in terms of econometric methods and data usage (for example, Currie and Thomas, 1995; Garces, Thomas, and Currie, 2002; Deming, 2009; Carneiro and Ginja, 2014; Johnson and Jackson, 2019; Ludwig and Miller, 2007; Bailey, Sun, and Timpe, 2021; Anders, Barr, and Smith, 2023). However, Head Start is inherently different from UPK, as it is a targeted program that provides a more comprehensive set of services.

[8] Berlinski, Galiani, and Manacorda (2008) also compare sibling outcomes in the context of a preschool expansion, using within-family variation to evaluate Uruguay's rollout of pre-primary education.

affected cohorts, while my design isolates the effect of starting pre-K at age three rather than at age four. Both parameters are of interest: The ITT reflects the effects of the program as a whole, including its spillovers onto schools and localities; the treatment effect measures what attendance itself contributes—and its heterogeneity opens a window onto the mechanisms at work.[9]

The remainder of this article is organized as follows. Section 2 provides background information on the UPK expansion and the Israeli context. Section 3 presents the data and sample construction. Section 4 outlines the empirical framework and validates its underlying assumptions. Section 5 presents the main analysis, examining how pre-K attendance at age three (versus age four) affects school progression and educational achievement. Section 6 explores heterogeneity in the results. Section 7 investigates potential mechanisms behind the gender differences in public pre-K returns. Section 8 concludes.

2 Background

Provision and expansion of public pre-K in Israel. Prior to 1999, local municipalities in Israel had the option to choose whether to provide pre-K education. These local programs were subject to regulation by the central government, which determined tuition fees, teacher requirements, and other regulations. Subsidies were available based on household income and were jointly funded by the central government and local municipalities, with subsidy rules established at the national level. This setup created an incentive for poorer municipalities, particularly those with a high proportion of low-income households—most of which were majority Arab—to refrain from opening pre-K programs (Kimhi, 2012).[10]

Israel launched its UPK program in 1999 by amending the Free and Compulsory Education Law to cover children as young as three years old, expanding beyond its original K–12 scope. The law outlined a 10-year rollout by locality, empowering the minister of education to determine

[9]These differences in estimands can also account for why the estimates I obtain are smaller than those in DS. First, the ITT estimates may capture indirect spillover effects of the expansion on non-participating peers, consistent with recent work showing that public pre-K can generate broader equilibrium effects (Berne, 2024; Jackson, Turner, and Bastian, 2025). Second, DS estimate the effect of a reform that expanded access to pre-K at ages three and four, whereas I estimate the effect of attending at age three; an age gradient in impacts is consistent with evidence from studies of childcare at ages zero to three documenting adverse effects on child development (Baker, Gruber, and Milligan, 2008; Fort, Ichino, and Zanella, 2020).

[10]Poor Jewish-majority municipalities were more likely to provide pre-K, as they received assistance from the central government through the Priority Localities program. The Supreme Court ruled in 2006 that the criteria for receiving this funding were discriminatory.

which areas would implement the reform. Using this authority, the minister prioritized specific municipalities and neighborhoods based on three criteria: municipalities in the bottom two deciles of an economic deprivation index, areas designated as confrontation lines, and regions classified as national priority areas. Since most Jewish localities already offered public pre-K before 1999, the pre-K expansion primarily benefited Arab children. The reform was implemented through two ministerial orders issued in 1999 and 2001, which specified the localities where the Free and Compulsory Education Law would apply to three- and four-year-olds. Two orders were necessary because many eligible municipalities were not initially equipped with the infrastructure needed to provide pre-K services to all children as required by law. Although the original plan was meant to be implemented nationwide, political and economic constraints delayed nationwide UPK coverage until 2015.

The introduction of UPK led to a significant increase in pre-K attendance among Arab children. Panel A of Figure 1 illustrates that from 2000 to 2005, the proportion of Arab students attending pre-K at age three doubled, from approximately 30% to about 60%. During this period, pre-K attendance among the Jewish population remained stable and high. The analysis in this paper is focused on children who resided in the Arab-majority localities included in the 2001 order since my data on student-level pre-K attendance start only in 2000. Panel B of Figure 1 shows that, similarly to the entire Arab population, Arab municipalities included in the 2001 order experienced a dramatic 20 percentage point increase in public pre-K attendance. Importantly, the baseline pre-K attendance is quite high (70%), as some municipalities included in the order were quicker in rolling out pre-K classrooms after the 1999 amendment.

Quality of Israeli public pre-K. In Israel, public pre-K programs operate for six hours a day, six days a week, 10 months a year. The Ministry of Education mandates that each class have a teacher and an assistant teacher. The teacher must hold a teaching certificate and, in many cases, also a bachelor's degree (Kimhi, 2012). The average class size in my sample was 32, resulting in a relatively low adult-to-child ratio of 1:16. For comparison, Norway, Germany, and Spain have adult-to-child ratios of 1:8, 1:12.5, and 1:13, respectively (Cornelissen et al., 2018; Havnes and Mogstad, 2011). While I lack data on expenditure per student for the period under study, in 2019, the estimated expenditure per student was $4,300 (PPP), significantly lower than that of most

European countries and Head Start.[11]

The Arab population in Israel. Israel's population includes a significant Arab minority. Arab citizens of Israel, also known as Palestinian citizens of Israel,[12] represent about 20% of all Israeli citizens and numbered 1.3 million people at the end of 2004. Arabs in Israel belong to three major religious groups: 83% are Muslims, 9% are Christians, and 8% are Druze (CBS, 2005b). The majority of the Arab population (80%) live in majority-Arab municipalities or villages (Haddad Haj-Yahya et al., 2021). There are substantial economic disparities between the Arab and Jewish populations in Israel. In 2004, income of Arab households was 36% lower than that of Jewish households (CBS, 2004). Part of this gap is driven by the lower participation of Arab women in the labor force, which in 2004 was 24% for women aged 25–64, compared with 79% for Jewish non-Ultraorthodox women (CBS, 2005a).

The low labor force participation rate of Arab women in Israel at that time was closely linked to childcare arrangements for young children. According to a 2004 survey by the Israeli Central Bureau of Statistics (CBS), 81% of Arab children aged 0–5 were cared for at home by a parent or unpaid relative.[13] This rate was significantly higher than in the Jewish population, in which only 30% of children in the same age group received such care. The disparity extended to private early-childhood programs as well: During the 2002/03 school year, just 0.9% of Arab children aged three attended private programs, compared with 4.6% of Jewish children.[14] Given these patterns, the counterfactual to public pre-K attendance was most likely home care by mothers or relatives.

The Israeli education system is largely separated, so Arab children typically attend Arabic-language public schools (from pre-K through high school) with predominantly other Arab students. Instruction in these schools is primarily in Arabic, with Hebrew taught as a second language and assessed within the national matriculation system; teachers in the Arab system are overwhelmingly Arab and largely female.

High school diploma (Bagrut). In Israel, high school students are examined in a series of

[11] In the same year, the estimated expenditure per student in Norway and Germany was $11,000 and $7,700, respectively. Source: `https://www.oecd.org/els/soc/PF3_1_Public_spending_on_childcare_and_early_education.pdf` (accessed 11/27/2024)

[12] This refers to Arab citizens within Israel's pre-1967 borders and does not include Palestinians residing in the West Bank or Gaza Strip.

[13] Source: `https://www.cbs.gov.il/he/publications/doclib/2004/social_survey/pdf/ty01.pdf` (accessed 11/27/2024)

[14] Source: `https://www.cbs.gov.il/he/publications/DocLib/2004/8.%20Shnaton%20Education/st08_06.pdf` (accessed 11/27/2024)

centrally administered matriculation exams. Each exam is associated with a number of credits (1 to 5), and to be awarded a high school diploma, students must receive a passing grade in exams collectively worth at least 20 credits. Some subjects (such as English and math) have a minimum credit requirement, and all students have to take at least the lowest-level exam in these subjects. I therefore define an advanced subject as one in which a student earns more than the minimum required credits. A Bagrut diploma is an important prerequisite for postsecondary education and a requirement for many entry-level positions in the labor market.

3 Data and Sample Construction

The data used in this study are sourced from administrative records of the Israeli Ministry of Education. I linked two types of files. First are enrollment records, which cover all students in the public education system. These records include details on each student's grade, school identifier, and demographic characteristics such as date of birth, gender, ethnicity, immigration status, and parental education. Additionally, these records contain unique identifiers for parents, enabling sibling identification. The coverage extends from the 1995/96 school year to 2018/19 for students in 1st through 12th grades, and from 1999/2000 to 2017/18 for public pre-K and kindergarten participants. The second type of file contains data on achievements on high school matriculation tests. These outcomes include an indicator for whether a diploma was awarded, total number of credits awarded, and indicators for advanced-level study in English, math, and Hebrew, which are important compulsory subjects.

I begin by constructing a census of all children born between 1996 and 2001.[15,16] These children were aged three during school years 1999/2000 through 2004/5, coinciding with the pre-K expansion in Arab municipalities. I then restrict my sample to Arab-majority municipalities that were included in the special order issued in 2001.[17] This sample restriction leaves 33 municipalities and an average of about 9,000 students per cohort, compared with an average of 32,520 Arab students per cohort in the population. Column 1 of Table 1 presents summary statistics for the entire analysis sample,

[15] I can observe nearly all children in these birth cohorts, as almost all children in Israel attend publicly funded schools and thus appear in my data when they enroll in first grade.

[16] The 1996 birth cohort is the first with available data on pre-K attendance at age three, while 2001 is the last cohort for which I can observe high school graduation outcomes.

[17] In the main analysis, I do not use municipalities included in the 1999 special order since my pre-K enrollment data only begin in 2000.

while column 2 focuses on children included in the family fixed-effects analysis—that is, children from families with at least two children in the main sample. The sibling sample is characterized by slightly larger families, lower parental education, and slightly lower academic achievements compared with the main sample.

4 Empirical Strategy

A simple comparison of students who attended pre-K with those who did not is potentially biased, as the decision to enroll a child is correlated with other significant determinants of human capital. Table 2 illustrates this. The table displays coefficients from regressing pre-K attendance on student and family-background characteristics. Column 1 presents results from univariate models, while columns 2 and 3 show results from a multivariate model that includes all the variables (column 2) and municipality-cohort fixed effects (column 3). The results reveal significant selection in who enrolls their children in public pre-K: Children of parents with higher education levels and children born earlier in the year are more likely to attend pre-K. As shown in Appendix Table A1, this selection into treatment biases OLS estimates of public pre-K's effect on educational outcomes: The pre-K coefficient decreases substantially when controls are added, suggesting substantial selection on both observable and potentially unobservable characteristics (Oster, 2019).

To overcome this selection bias, I compare siblings within the same family who attended and did not attend public pre-K by estimating a family fixed-effects model. The family fixed-effects approach controls for family unobserved characteristics that are time invariant. This method has been used in the preschool literature, mainly in the context of Head Start (Currie and Thomas, 1995; Garces, Thomas, and Currie, 2002; Deming, 2009). Formally, I estimate the following model:

$$y_i = \alpha + \beta PreK_i + X_i\gamma + \delta_{j(i)} + \tau_{m(i),c(i)} + \epsilon_i \tag{1}$$

Here, y_i is child i's educational outcome, $PreK_i$ is an indicator for whether child i attended public pre-K at age three, X_i is a vector of child i's background characteristics (which include sex, month of birth, and birth order), $\delta_{j(i)}$ is a family fixed effect, and $\tau_{m(i),c(i)}$ is a municipality-by-birth-cohort fixed effect.

The primary focus of this paper is β. For β to have a causal interpretation as the effect of attending pre-K, the identifying assumption requires that within a family, attendance in a public pre-K program is not correlated with a child's unobserved characteristics. Indeed, column 4 of Table 2 shows that within-family variation in pre-K enrollment is not correlated with predetermined characteristics such as sex or birth order, though, as discussed later, it is still correlated with month of birth, which I control for flexibly. Estimates of β could be biased if, for instance, parents choose to enroll children with lower endowments in public pre-K as a compensatory measure; this could lead to an underestimate of the true effect. Conversely, if mothers who lose their jobs decide against sending their children to public pre-K, this could result in an overestimate of the true effect, as pre-K attendance will be positively correlated with household income at age three. Considering that the program is free and demands minimal investment from parents, both scenarios are unlikely.

The most probable cause of within-family variation in pre-K attendance is the substantial expansion of public pre-K in Arab municipalities during the sample period. Crucially, the variation it induces is plausibly uncorrelated with unobserved differences among siblings. To further investigate variation in public pre-K attendance within families, I adapt concepts from the local average treatment effect framework (Imbens and Angrist, 1994), with a slight abuse of terms. I categorize families into four types based on the pre-K enrollment status at age three of the oldest and youngest child:[18] "never takers" (who sent none of their children to public pre-K), "always takers" (who sent all their children), "expansion compliers" (who sent their youngest but not eldest children), and "expansion defiers" (who sent their eldest but not youngest child). Following Miller, Shenhav, and Grosz (2023), I term the expansion compliers and defiers collectively as "switchers." Importantly, only switchers help identify β (Miller, Shenhav, and Grosz, 2023).

The pattern that emerges in Table 3 is consistent with the idea that within-family differences in pre-K attendance are due to the expansion: Expansion compliers—families who, in line with increased availability of pre-K, send only younger siblings—make up about 80% of switchers. Expansion defiers are few—727 families, about 4% of families in the sibling sample—and their enrollment pattern is consistent with rationing by birth timing rather than selection on child-specific endowments: Their youngest children are disproportionately born late in the year (50.6% were born September–December, compared with 35.3% among expansion compliers). Such a pattern

[18]The classification is based on the oldest and youngest siblings present in the sibling sample.

would be expected if municipalities prioritized older children when slots were scarce—a channel that operates primarily through month of birth.

The expansion is not the only supply-side source of within-family variation. Capacity constraints persisted throughout the study period: The reform required several years to become truly universal because many municipalities lacked the infrastructure to serve all eligible children. The idea that rationing in a free, capacity-constrained program generates within-family variation in attendance that is plausibly orthogonal to differences in children's endowments is not new and has been invoked in the context of Head Start (e.g., Deming, 2009). The data are consistent with this account: within-family variation in enrollment is uncorrelated with observed proxies for child endowments and household circumstances at age three (as I show below using supplemental data), and its only observed predictor is month of birth—itself in line with the prioritization of older children when slots were scarce. To address any remaining concern, I also show below that the results are robust to excluding expansion defiers altogether.

The switcher analysis suggests that there are two important variables to control for. First, since I mostly compare older siblings to younger ones, a cohort fixed effect is necessary; otherwise, the estimates might conflate time trends with the treatment effect. Allowing these trends to vary by municipality, the municipality-by-birth-cohort fixed effect $\tau_{m(i),c(i)}$ additionally absorbs any common shock within a cell, such as peer effects from more pre-K-treated classmates. Second, as discussed above, month of birth predicts within-family variation in enrollment, which is why equation 1 includes month-of-birth fixed effects. Importantly, both sets of parameters are identified off variation among siblings in always-taker and never-taker families.

The switcher analysis also reveals that most children who do not attend pre-K at age three do attend at age four. This can be seen in the last two rows of Table 3, which display pre-K attendance rates at ages three and four, respectively, for the oldest sibling across the four family types. By definition, the share of oldest children attending pre-K at age three is zero for never-taker and expansion-complier families, and one for always-taker and expansion-defier families. Among expansion-complier families, 73% of oldest siblings attended pre-K at age four. Therefore, my identification strategy primarily compares older children who did not attend pre-K at age three but did attend at age four with their younger siblings who started attending at age three. Thus, I interpret the results as the effect of beginning pre-K one year earlier—at age three rather than

four.

5 Results

Table 4 presents results from estimating equation 1 for various outcomes. Panel A focuses on school-progression outcomes. The estimates suggest that pre-K attendance significantly enhances students' progression through the school system. Columns 1–3 show that children who attended pre-K are 1.9 percentage points more likely to start primary school on time, 1.3 percentage points less likely to repeat a grade, and 0.9 percentage points more likely to advance to 11th grade. While the second and third estimates are only marginally statistically significant, column 4 shows that all three factors contribute to a statistically significant improvement in on-time high school graduation by 2.4 percentage points. However, this effect primarily reflects better progression rather than improved educational attainment, as the impact on overall high school graduation is smaller and not statistically significant (column 5).

Panel B further indicates that pre-K at age three plays a limited role in improving educational achievements. Column 1 shows a small and statistically nonsignificant effect on obtaining a Bagrut diploma. While column 2 suggests some improvement in the quality of the diploma, as measured by an increase in diploma credits, this increase does not stem from taking additional credits in key subjects such as mathematics, English, and Hebrew (columns 3–5).

In the main analysis in Table 4 I include expansion defiers in the sample for two reasons. First, as discussed above, the decision not to send the youngest child is plausibly exogenous, potentially stemming from local supply conditions at the time the child was about to enter pre-K; consistent with this, these children are disproportionately born late in the year (and I control for month of birth). Second, the division into four family types is somewhat simplistic since families may have more than two children in the sample. However, Table 5 demonstrates that my results are robust to excluding expansion defiers. The findings are also robust to several additional specification checks. First, Table 5 shows that the results hold when imposing common cohort fixed effects across all municipalities, rather than municipality-specific cohort fixed effects. Second, the results are not sensitive to controlling for month of birth. Third, while my main analysis focuses on Arab-majority municipalities included in the 2001 special order, expanding the sample to include all Arab-majority

municipalities reveals similar patterns.

Finally, to address concerns about families with more than two children receiving higher weight (Miller, Shenhav, and Grosz, 2023), I estimate models using only the oldest and youngest child from each family in the family fixed-effects sample. These estimates remain consistent with the main analysis. Keeping two siblings per family fixes family size and equalizes the variation in treatment across all switcher families—two factors that Miller, Shenhav, and Grosz (2023) show can lead to a biased estimate of the treatment effect for the target population. However, this approach has the drawback of reducing the sample size, resulting in less precise estimates. In Appendix Table A2, I implement the reweighting procedure introduced by Miller, Shenhav, and Grosz (2023). Intuitively, this method downweights families that are overrepresented in the switcher group and upweights those that more closely resemble pre-K participants. The resulting estimates closely mirror the main results, though most are somewhat larger in magnitude.

To contextualize my results, I compare my estimates to existing findings. I begin by examining estimates from large-scale pre-K programs that report treatment effects on similar outcomes and provide standard errors or confidence intervals. Panel A of Figure 2 shows that my estimates are smaller than other studies', though they generally fall within these studies' confidence intervals.[19] The weaker effects of the Israeli expansion may stem from the program's lower quality as reflected in its lower per-student expenditure. Notably, despite the smaller effects, the results demonstrate that even lower-quality programs can improve school-progression outcomes, leading to more on-time high school graduation and potentially earlier labor force entry.

Next, I compare my findings to estimates from DS, who also study pre-K expansion in Arab municipalities in Israel.[20] Panel B of Figure 2 shows that my estimates are smaller than those in DS for overlapping outcomes. Two differences between the estimands can account for this gap. First, if the expansion affects non-enrollees, for example through changes in peer interactions or broader impacts on the local economy (Jackson, Turner, and Bastian, 2025; Berne, 2024), then the ITT reflects both direct effects on participants and spillovers onto non-participants. In that case, the

[19] A notable exception is Berlinski, Galiani, and Manacorda (2008), who find substantial improvements in school progression by age 15, the oldest age they observe, whereas I find a small and statistically insignificant effect of pre-K on the probability of ever attending 11th grade.

[20] To enable appropriate comparison, I derive treatment effect estimates by dividing the ITT estimates by the first stage reported in DS. I construct confidence intervals by similarly scaling the implied confidence intervals from their study using the ITTs' standard errors.

Wald ratio (ITT divided by take-up) need not identify the effect of attendance and may overstate it. Second, I estimate the effect of starting pre-K at age three rather than four, while DS study an expansion that increased enrollment at both ages three and four (Figure 2 in DS). This gradient in effect sizes, in which impacts increase with child age, aligns with evidence from studies of public childcare for children aged zero to three that document negative impacts on child development (Baker, Gruber, and Milligan, 2008; Fort, Ichino, and Zanella, 2020). While a few studies examine pre-K at age three (Blanden et al., 2016; Cornelissen et al., 2018; Felfe, Nollenberger, and Rodríguez-Planas, 2015), the majority of the literature focuses on age four (or ages three and four combined), and little is known about how returns to pre-K at age three compare to returns at age four in the same context. Taken together, this study and DS suggest that expanding pre-K at age four yields higher returns.

5.1 Threats to Identification

The causal interpretation of β in equation 1 relies on two key assumptions: (1) within-family variation in pre-K attendance is uncorrelated with unobserved individual-level factors, and (2) there are no spillover effects on siblings from attending pre-K. While I cannot directly verify these assumptions, in this subsection I present diagnostic tests that provide supporting evidence for them.

5.1.1 within-family Selection

If parents systematically choose to send one child but not another to pre-K based on child-specific traits or household conditions at the time the child becomes eligible for enrollment, this could introduce bias into the model estimates. In the context of early childhood, one concern is that the allocation of family resources involved in this decision may depend on children's endowments. While such selection is plausible when pre-K attendance involves financial costs, it is less clear why parents would differentially enroll siblings when pre-K is free and requires minimal investment from parents. Empirically, column 4 of Table 2 shows no association between within-family variation in pre-K enrollment and either child sex or birth order. The only within-family predictor is month of birth, which, as discussed earlier, I control for throughout.

I further provide evidence on the association between birth weight and pre-K enrollment. Birth weight is often used in the literature as a proxy for a child's endowment and is a predictor and

determinant of later-life outcomes (Black, Devereux, and Salvanes, 2007). Thus, such analysis can indicate whether parents selectively send one sibling and not the other to pre-K. To perform this analysis, I use supplemental de-identified individual-level records from the Israeli Central Bureau of Statistics that include child birth weight and enrollment information. These data cannot be matched to the main education administrative dataset and are available only from 2006 onward, so they do not cover the period of the main analysis sample. In Panel A of Appendix Table A4, column 2 shows that, in the cross-section, birth weight is positively correlated with pre-K enrollment. However, once I include family fixed effects in column 3, within-family differences in birth weight no longer predict differences in pre-K enrollment across siblings.

Another concern is that variation in household circumstances at the time of pre-K eligibility could generate within-family differences in enrollment. For instance, mothers who do not send their child to pre-K might reduce their labor supply, leading to differences in household resources at age three between siblings and thereby confounding the causal interpretation of the estimates. To explore this possibility, I also use the supplemental data. In Panel B of Appendix Table A4, column 2 shows that, in the cross-section, pre-K enrollment is positively associated with maternal employment. However, once I include family fixed effects in column 3, this relationship disappears. Taken together, these tests suggest that within-family differences in pre-K attendance are not systematically driven by either child endowments or maternal employment at age three.

5.1.2 Sibling Spillovers

To examine whether one sibling's pre-K attendance affects another's educational outcomes, I conduct two analyses. Both are based on the idea that spillover effects should increase with the degree of siblings' exposure to one another (Black et al., 2021).

In the first analysis, I exploit the fact that siblings who are closer in age are more likely to spend time together and, consequently, may exhibit stronger spillover effects from a sibling's pre-K attendance. Specifically, I restrict the sample to the eldest and youngest child in each household (a sample for which I have already shown in Table 5 that the main results remain unchanged) and augment equation 1 with an interaction term between pre-K attendance and the age spacing between the two siblings.

If positive within-family spillover effects exist and diminish with age spacing, the interaction

term should be positive, as the estimated effect of β would be more diluted when siblings are close in age. However, as shown in Panel A of Appendix Table A5, the estimated interaction is not statistically different from zero, whether I specify a linear interaction (column 1) or a non-linear spacing dummy (column 2). These results are, of course, only suggestive, since sibling spacing is correlated with other important family characteristics.

My second analysis reproduces a difference-in-differences model introduced in Black et al. (2021) to capture sibling spillover effects. In a nutshell, under the assumption that time spent together increases spillovers, a second-born child should be more affected than a first-born by shocks to a third-born sibling. In my setting, the relevant "shock" is whether the third-born attended pre-K.

Accordingly, I restrict my sample to families with three children, keeping only the first- and second-born, and estimate the following difference-in-differences specification:

$$y_{ij} = Family_j + \lambda Child2_i + \theta(Child2_i \times Treated_j) + X_{ij}\gamma + \epsilon_{ij} \tag{2}$$

where y_{ij} is an educational outcome for child i in family j, $Family_j$ denotes a family fixed effect, and $Child2_i$ is an indicator for the second-born child. $Treated_j$ is a family-level indicator for whether the third-born child attended pre-K. The coefficient of interest, θ, captures the spillover effect under the assumptions described above.

Column 1 of Panel B in Appendix Table A5 reports the estimate of θ for the analysis sample. The coefficient is small and statistically insignificant, although imprecisely estimated. To improve precision, I re-estimate equation 2 using the full sample of children from Arab-majority municipalities. The estimate remains small but becomes more precisely estimated. I conclude that sibling spillovers are unlikely to bias the estimates from the family fixed-effects model.[21]

6 Heterogeneous Effects

I next analyze how effects vary across student subgroups. While prior research often reveals heterogeneous impacts of public pre-K, most long-term studies can only estimate ITT effects because data on program enrollment are incomplete. By contrast, I am able to estimate treatment effects

[21] This finding is consistent with DeMalach and Schlosser (2026), showing that an earlier pre-K expansion in northern Arab municipalities in Israel did not affect maternal labor supply, a key channel through which sibling spillovers from pre-K might operate.

directly, providing new insights into how the impact of public pre-K varies across populations.

Table 6 reports estimates of the fixed-effects model for the school-progression outcomes for different subgroups.[22] Since this exercise involves cutting the sample multiple times into smaller subsamples, I report for each sample cut the p-value from a test that all differences are zero for all outcomes.

Columns 1 and 2 present results for boys and girls, respectively. The point estimates are consistently larger for girls: Girls who attend pre-K are more likely to start first grade on time than boys, are less likely to repeat a grade, and show a more pronounced increase in the probability of reaching 11th grade. These differences translate to a larger effect on girls' on-time high school graduation rates compared with boys. Notably, I find a significant effect on high school completion for girls—a finding not observed in the pooled sample. I can reject the null hypothesis of no differences across all outcomes with high statistical significance. In the next section, I explore what may explain this striking difference in treatment effects between boys and girls.

Columns 3 and 4 examine heterogeneity by relative age, comparing students born earlier versus later in the birth year.[23] This analysis is particularly relevant because month of birth is a strong predictor of enrollment: Children born later in the year are less likely to attend public pre-K (Table 2). The point estimates suggest larger benefits for younger students, especially for starting first grade on time and graduating high school on time, and a joint test rejects equality of effects across outcomes. These patterns suggest reverse selection on gains, whereby the group that appears to benefit more (younger children) is less likely to receive treatment.

I also examine heterogeneity by maternal education and family size.[24] While maternal education strongly predicts pre-K attendance (Table 2), columns 5 and 6 of Table 6 show it also predicts gains from pre-K: Children of more-educated mothers benefit more on average.

Columns 7 and 8 reveal no statistically significant differences by family size, though the point estimates indicate potentially larger benefits for children from larger households.

[22] Appendix Table A3 shows heterogeneity analysis for educational achievements, for which, as in the main analysis, I do not find an effect.

[23] I define older students as those born from January through August, and younger students as those born from September to December. This choice aligns with the school-entry-age cutoff date and the Ministry of Education policy that considers the latter group less ready to start first grade on time, thus creating more lenient conditions for redshirting younger children.

[24] Maternal education is missing for 18.5% of students. The main results are robust to restricting the sample to students with non-missing maternal education.

7 Mechanisms Behind Heterogeneity by Sex

What mechanisms might explain the striking gender differences in returns to UPK? A hypothesis can be formed based on the voluminous literature on the early origins of gender gaps in educational achievements (Reardon et al., 2019; Autor et al., 2016; Lavy and Sand, 2018). This literature finds that girls have early advantages in social and behavioral skills, which may enhance their ability to benefit from pre-K environments. Specifically, boys are more prone to experiencing temperament and self-regulation challenges (Bertrand and Pan, 2013; DiPrete and Jennings, 2012), whereas girls exhibit a greater propensity for engaging in constructive interactions with both peers and adults (Magnuson et al., 2016). Furthermore, girls tend to derive more benefit from unstructured curricula (Fidjeland et al., 2023). These developmental disparities are reflected in school-readiness assessments, where boys are more frequently perceived as less prepared for first grade and consequently experience higher rates of redshirting (Cook and Kang, 2020; Deming and Dynarski, 2008). This apparent lag in boys' noncognitive skill development has led some experts to propose extending their kindergarten experience by an additional year (Reeves, 2022).

Thus, it is possible that the pre-K classroom environment is more conducive to girls' learning because of their better interactions with teachers, greater benefits from unstructured play, and overall higher level of readiness. Of these three factors, the readiness hypothesis yields a testable prediction in my data. This hypothesis posits that boys' skills develop at a slower rate than girls', so that boys arrive at pre-K less ready to benefit from the classroom environment. If differential school readiness were driving the heterogeneity in returns to UPK, young boys, presumed to be the least ready, should benefit the least from attending UPK. In Panel A of Figure 3, I test whether sex differences in readiness explain the gender gap in UPK returns by estimating heterogeneous effects of UPK attendance on graduating from high school on time—which, as shown in the previous section, serves as a good summary index for overall school progression.[25] The pattern that emerges in the figure contradicts the readiness hypothesis: The point estimate for young boys is positive and statistically larger than the estimate for older boys.

What else can explain heterogeneity in the returns to UPK? The literature on early childhood education highlights the importance of the quality of the counterfactual care environment in gener-

[25] In Appendix Figures A1 to A4, I show that these patterns are largely consistent across other school-progression outcomes.

ating treatment heterogeneity (Kline and Walters, 2016). Since most children in this setting would remain in home care in the absence of UPK, sex differences in the home environment emerge as a potential explanation for treatment heterogeneity. In this setting—Arab municipalities with a Muslim majority—gender norms and parental preferences for boys may create different home environments for young children: Girls typically stay home with a parent or other adult relatives and participate in household chores, while boys are more likely to spend time outdoors and play with peers (UNDP, 2006; El Feki, Heilman, and Barker, 2017). Consequently, the introduction of UPK may have represented a more dramatic change for girls by providing substantial exposure to peer play and nonrelative-adult interactions. For boys, these experiences were already more accessible even in the absence of UPK. Moreover, parental investments could play a crucial role. Given that favoritism toward boys is well documented in Arab and Muslim communities (Yassine-Hamdan and Strate, 2020), the introduction of UPK might have provided girls with access to quality time with adults, while for boys, the change in adult interaction was less substantial.

I provide two tests for the home-care-quality hypothesis. First, I examine whether the heterogeneity by sex is concentrated in more traditional households, using maternal education as a proxy. Since gender norms and son preference are expected to be more pronounced in more traditional households, under the home-care-quality hypothesis the heterogeneous effect across sexes should be driven by this group. Panel B of Figure 3 confirms this prediction: Girls from more traditional households benefit significantly more than boys from pre-K attendance, while in less traditional households there is no girl advantage and, if anything, the point estimate favors boys.

Second, I examine how the girl advantage varies with sibling-sex composition, comparing families with both sons and daughters to families with children of only one sex. Since there should be less scope for son preference in investments in families with only daughters, the home-care-quality hypothesis predicts that the heterogeneous effect should be driven by girls who have brothers. Panel C of Figure 3 corroborates this prediction: Girls from mixed-sex families gain significantly more from attending pre-K than boys, while in single-sex families there is no statistical difference between girls and boys.

The analysis in this section indicates that the sex differences in returns to pre-K are more consistent with differences in home-care environments between girls and boys than with differences in developmental readiness to benefit from pre-K. As in most heterogeneity-based mechanism analyses,

a caveat applies: month of birth, maternal education, and sibling-sex composition are potentially correlated with other dimensions along which effects may vary. That said, any competing explanation must account not only for heterogeneity along these dimensions, but also for its interaction with sex.

8 Conclusions

This paper provides estimates of the long-term treatment effects of starting pre-K at age three rather than at age four. Leveraging comprehensive administrative data from Israel and within-family variation in pre-K attendance generated largely by supply-side conditions, I estimate the effects of attending pre-K at age three on school progression and educational achievement at the end of high school. I find that attending pre-K at age three significantly improves school progression but has limited effects on academic achievements. These findings suggest that while earlier pre-K entry can help children progress through school more smoothly, it might not substantially enhance their academic performance. While, in theory, smoother school progression can lead to welfare gains and increased public revenues—as children enter the labor force earlier and begin earning and paying income taxes sooner—in practice, under a wide set of assumptions, these benefits are small and insufficient to justify pre-K at age three on their own. Applying the marginal value of public funds (MVPF) framework proposed by Hendren and Sprung-Keyser (2020), I calculate an MVPF that is substantially below the benchmark value of 1 for a non-distortionary cash transfer.[26] This calculation likely understates the true return to pre-K for two reasons. First, it does not incorporate potential effects on adult earnings, which may be positive given the improvements in school progression even if high school achievements are unaffected. Second, it omits indirect effects of the expansion on non-participants and the local economy (Berne, 2024; Jackson, Turner, and Bastian, 2025).

The analysis reveals striking heterogeneity in treatment effects across sexes: Point estimates

[26] For example, assuming that children enter the labor force right after graduating from high school and earn NIS 7,000 in the first year with a 4% annual increase, children who graduate on time start earning at age 19 (16 years after pre-K) and children who graduate one year later start earning at age 20 (17 years after pre-K). The difference in discounted lifetime income between these two cases is NIS 14,700, where $t = 0$ is the year of pre-K attendance and the discount rate is 3%. Assuming a 20% income tax and an effect of 0.024 on graduating on time, I calculate that the government saves about NIS 70 for every child attending pre-K. This is significantly lower than the expenditure per child of about NIS 16,000 in 2019. The MVPF is thus: $\frac{0.024 \times 0.8 \times 14{,}700}{16{,}000 - 0.024 \times 0.2 \times 14{,}700} \approx 0.018$. Incorporating fiscal savings from reduced grade repetition does not meaningfully affect this calculation.

favor girls on every progression outcome, and a joint test decisively rejects equal effects for boys and girls. I test two explanations for this pattern: differential developmental readiness and differences in the counterfactual home environment. The evidence is inconsistent with the readiness explanation—relatively younger boys, presumed to be the least ready, benefit more rather than less—and more consistent with the home-care explanation: A significant girl advantage appears in families where gender norms and son preference are more likely to operate (those with low maternal education and those raising both boys and girls) but not in their counterparts. The rejection of the readiness explanation also speaks to proposals to give boys more time to mature before formal schooling (Reeves, 2022): In this setting, the boys presumed least ready benefit the most from starting a year earlier. More broadly, these findings highlight the potential for UPK to serve as an equalizing force by providing opportunities that some children, particularly girls in traditional settings, might not otherwise receive at home. While this is a desirable outcome, it will contribute to widening the gender educational gap, as girls outperform boys across all educational outcomes in this setting.

The findings of this paper raise several questions for future research. First, are the smaller effects I find (compared with other studies' findings) the result of estimating the impact of attending pre-K at age three rather than at age four? Or are they the result of a lower-quality pre-K program, as reflected by the lower expenditure per child, larger classes, and lower teacher qualifications? Alternatively, the gap may reflect the difference between treatment effects and ITT estimates: since most existing long-term studies estimate ITT effects, their larger magnitudes may partly capture indirect effects of pre-K expansions on non-participants—suggesting that such spillovers are a quantitatively important component of what universal pre-K does. In addition, as I stop the analysis at the end of high school, it remains unknown whether later-life outcomes will follow the pattern of school-progression outcomes (which show significant improvements) or educational-achievement outcomes (which show limited effects). Moreover, a crucial question is whether it is possible to modify the preschool program in ways that maintain the benefits for girls while also generating positive effects for boys. Finally, the home-counterfactual account offers a way to read the conflicting cross-country evidence on which sex gains more from universal pre-K: The same program may produce different gender gaps depending on the home environments it replaces.

Declaration of generative AI and AI-assisted technologies in the manuscript preparation process

During the preparation of this work the author used Claude (Anthropic) in order to assist with editing and revising the text. After using this tool, the author reviewed and edited the content as needed and takes full responsibility for the content of the published article.

References

Aizer, Anna, Paul Devereux, and Kjell Salvanes. 2022. "Grandparents, Moms, or Dads? Why Children of Teen Mothers Do Worse in Life." *Journal of Human Resources* 57 (6): 2012–2047.

Akee, Randall, and Leah R Clark. 2024. "Preschool Lottery Admissions and Its Effects on Long-Run Earnings and Outcomes." Working Paper 32570, National Bureau of Economic Research.

Almond, Douglas, Janet Currie, and Valentina Duque. 2018. "Childhood Circumstances and Adult Outcomes: Act II." *Journal of Economic Literature* 56 (4): 1360–1446.

Anders, John, Andrew C. Barr, and Alexander A. Smith. 2023. "The Effect of Early Childhood Education on Adult Criminality: Evidence from the 1960s through 1990s." *American Economic Journal: Economic Policy* 15 (1): 37–69.

Attar, Itay, and Danny Cohen-Zada. 2018. "The Effect of School Entrance Age on Educational Outcomes: Evidence Using Multiple Cutoff Dates and Exact Date of Birth." *Journal of Economic Behavior & Organization* 153: 38–57.

Autor, David, David Figlio, Krzysztof Karbownik, Jeffrey Roth, and Melanie Wasserman. 2016. "School Quality and the Gender Gap in Educational Achievement." *American Economic Review* 106 (5): 289–295.

Bailey, Martha J., Shuqiao Sun, and Brenden Timpe. 2021. "Prep School for Poor Kids: The Long-Run Impacts of Head Start on Human Capital and Economic Self-Sufficiency." *American Economic Review* 111 (12): 3963–4001.

Baker, Michael, Jonathan Gruber, and Kevin Milligan. 2008. "Universal Child Care, Maternal Labor Supply, and Family Well-Being." *Journal of Political Economy* 116 (4): 709–745.

Baker, Michael, Jonathan Gruber, and Kevin Milligan. 2019. "The Long-Run Impacts of a Universal Child Care Program." *American Economic Journal: Economic Policy* 11 (3): 1–26.

Baker, Michael, and Kevin Milligan. 2016. "Boy-Girl Differences in Parental Time Investments: Evidence from Three Countries." *Journal of Human Capital* 10 (4): 399–441.

Barcellos, Silvia Helena, Leandro S. Carvalho, and Adriana Lleras-Muney. 2014. "Child Gender and Parental Investments in India: Are Boys and Girls Treated Differently?" *American Economic Journal: Applied Economics* 6 (1): 157–189.

Berlinski, Samuel, Guillermo Cruces, Sebastian Galiani, Paul Gertler, and Fabian Gonzalez. 2025. "Long-run Effects of Universal Pre-Primary Education Expansion: Evidence from Argentina." Working Paper 34552, National Bureau of Economic Research.

Berlinski, Samuel, Sebastian Galiani, and Paul Gertler. 2009. "The effect of pre-primary education on primary school performance." *Journal of Public Economics* 93 (1): 219–234.

Berlinski, Samuel, Sebastian Galiani, and Marco Manacorda. 2008. "Giving Children a Better Start: Preschool Attendance and School-Age Profiles." *Journal of Public Economics* 92 (5): 1416–1440.

Berne, Jordan S. 2024. "The Long-Run Impacts of Universal Pre-K with Equilibrium Considerations." EdWorkingPaper 22-626, Annenberg Institute at Brown University.

Bertoni, Marco, Giorgio Brunello, and Lorenzo Cappellari. 2020. "Who benefits from privileged peers? Evidence from siblings in schools." *Journal of Applied Econometrics* 35 (7): 893–916.

Bertrand, Marianne, and Jessica Pan. 2013. "The Trouble with Boys: Social Influences and the Gender Gap in Disruptive Behavior." *American Economic Journal: Applied Economics* 5 (1): 32–64.

Black, Sandra E, Sanni Breining, David N Figlio, Jonathan Guryan, Krzysztof Karbownik, Helena Skyt Nielsen, Jeffrey Roth, and Marianne Simonsen. 2021. "Sibling spillovers." *The Economic Journal* 131 (633): 101–128.

Black, Sandra E., Paul J. Devereux, and Kjell G. Salvanes. 2007. "From the Cradle to the Labor Market? The Effect of Birth Weight on Adult Outcomes*." *The Quarterly Journal of Economics* 122 (1): 409–439.

Blanden, Jo, Emilia Del Bono, Kirstine Hansen, and Birgitta Rabe. 2022. "Quantity and quality of childcare and children's educational outcomes." *Journal of Population Economics* 35 (2): 785–828.

Blanden, Jo, Emilia Del Bono, Sandra McNally, and Birgitta Rabe. 2016. "Universal Pre-school Education: The Case of Public Funding with Private Provision." *The Economic Journal* 126 (592): 682–723.

Blau, Francine D., Lawrence M. Kahn, Peter Brummund, Jason Cook, and Miriam Larson-Koester. 2020. "Is there still son preference in the United States?" *Journal of Population Economics* 33 (3): 709–750.

Boxberger, Karolina, and Anne Kerstin Reimers. 2019. "Parental Correlates of Outdoor Play in Boys and Girls Aged 0 to 12—A Systematic Review." *International Journal of Environmental Research and Public Health* 16 (2): 190.

Bruhn, Jesse, and Emily Emick. 2023. "Lottery Evidence on the Impact of Preschool in the United States: A Review and Meta-Analysis." Discussion Paper 2023.20, Blueprints Labs.

Carneiro, Pedro, and Rita Ginja. 2014. "Long-Term Impacts of Compensatory Preschool on Health and Behavior: Evidence from Head Start." *American Economic Journal: Economic Policy* 6 (4): 135–173.

Cascio, Elizabeth U. 2021. "Early Childhood Education in the United States: What, When, Where, Who, How, and Why." NBER Working Papers 28722, National Bureau of Economic Research, Inc.

Cascio, Elizabeth U. 2023. "Does Universal Preschool Hit the Target?: Program Access and Preschool Impacts." *Journal of Human Resources* 58 (1): 1–42.

CBS. 2004. "Income Survey 2004.", Israel Central Bureau of Statistics. Jerusalem.

CBS. 2005a. "Population Aged 15 and Over and Population aged 25-54 (at main working ages), by Civilian Labour Force Characteristics, Population Group and Sex." Statistical Report 12.1, Israel Central Bureau of Statistics. Jerusalem.

CBS. 2005b. "Population Estimates: Population, by Religion and Population Group." Statistical Report 2.1, Israel Central Bureau of Statistics. Jerusalem.

Chetty, Raj, and Nathaniel Hendren. 2018. "The Impacts of Neighborhoods on Intergenerational Mobility I: Childhood Exposure Effects." *The Quarterly Journal of Economics* 133 (3): 1107–1162.

Cook, Philip J., and Songman Kang. 2020. "Girls to the Front: How Redshirting and Test-score Gaps are Affected by a Change in the School-entry Cut Date." *Economics of Education Review* 76: 101968.

Cornelissen, Thomas, Christian Dustmann, Anna Raute, and Uta Schönberg. 2018. "Who Benefits from Universal Child Care? Estimating Marginal Returns to Early Child Care Attendance." *Journal of Political Economy* 126 (6): 2356–2409.

Crockenberg, Susan C. 2003. "Rescuing the Baby From the Bathwater: How Gender and Temperament (May) Influence How Child Care Affects Child Development." *Child Development* 74 (4): 1034–1038.

Currie, Janet, and Duncan Thomas. 1995. "Does Head Start Make a Difference?" *The American Economic Review* 85 (3): 341–364.

Dahl, Gordon B., and Enrico Moretti. 2008. "The Demand for Sons." *Review of Economic Studies* 75 (4): 1085–1120.

DeMalach, Elad, and Analia Schlosser. 2026. "Short- and Long-Term Effects of Universal Preschool: Evidence from the Arab Population in Israel." *American Economic Journal: Economic Policy* 18 (2): 404–440.

Deming, David. 2009. "Early Childhood Intervention and Life-Cycle Skill Development: Evidence from Head Start." *American Economic Journal: Applied Economics* 1 (3): 111–134.

Deming, David, and Susan Dynarski. 2008. "The Lengthening of Childhood." *Journal of Economic Perspectives* 22 (3): 71–92.

DiPrete, Thomas A., and Jennifer L. Jennings. 2012. "Social and Behavioral Skills and the Gender Gap in Early Educational Achievement." *Social Science Research* 41 (1): 1–15.

Dossi, Gaia, David Figlio, Paola Giuliano, and Paola Sapienza. 2021. "Born in the family: Preferences for boys and the gender gap in math." *Journal of Economic Behavior & Organization* 183: 175–188.

El Feki, Shereen, Brian Heilman, and Gary Barker. edited 2017. *Understanding masculinities: results from the International Men and Gender Equality Survey (IMAGES) - Middle East and North Africa ; Egypt, Lebanon, Morocco, and Palestine*. Cairo and Washington, DC: UN Women and Promundo-US.

Else-Quest, Nicole M., Janet Shibley Hyde, H. Hill Goldsmith, and Carol A. Van Hulle. 2006. "Gender differences in temperament: A meta-analysis." *Psychological Bulletin* 132 (1): 33–72.

Felfe, Christina, Natalia Nollenberger, and Núria Rodríguez-Planas. 2015. "Can't Buy Mommy's Love? Universal Childcare and Children's Long-term Cognitive Development." *Journal of Population Economics* 28 (2): 393–422.

Fidjeland, Andreas, Mari Rege, Ingeborg F. Solli, and Ingunn Størksen. 2023. "Reducing the Gender Gap in Early Learning: Evidence From a Field Experiment in Norwegian Preschools." *European Economic Review* 154: 104413.

Figlio, David, Paola Giuliano, Riccardo Marchingiglio, Umut Ozek, and Paola Sapienza. 2024. "Diversity in Schools: Immigrants and the Educational Performance of U.S.-Born Students." *Review of Economic Studies* 91 (2): 972–1006.

Fitzpatrick, Maria D. 2008. "Starting School at Four: The Effect of Universal Pre-Kindergarten on Children's Academic Achievement." *The B.E. Journal of Economic Analysis & Policy* 8 (1).

Fort, Margherita, Andrea Ichino, and Giulio Zanella. 2020. "Cognitive and Noncognitive Costs of Day Care at Age 0–2 for Children in Advantaged Families." *Journal of Political Economy* 128 (1): 158–205.

Garces, Eliana, Duncan Thomas, and Janet Currie. 2002. "Longer-Term Effects of Head Start." *The American Economic Review* 92 (4): 999–1012.

Gazze, Ludovica, Claudia Persico, and Sandra Spirovska. 2024. "The Long-Run Spillover Effects of Pollution: How Exposure to Lead Affects Everyone in the Classroom." *Journal of Labor Economics* 42 (2): 357–394.

Gormley, William T., and Ted Gayer. 2005. "Promoting School Readiness in Oklahoma: An Evaluation of Tulsa's Pre-K Program." *The Journal of Human Resources* 40 (3): 533–558.

Gray-Lobe, Guthrie, Parag A Pathak, and Christopher R Walters. 2023. "The Long-Term Effects of Universal Preschool in Boston." *The Quarterly Journal of Economics* 138 (1): 363–411.

Haddad Haj-Yahya, Nasreen, Muhammad Khlaily, Arik Rudnitzky, and Ben Fargeon. 2021. "The Arab Society in Israel Yearbook: 2021." Annual Report, The Israel Democracy Institute. Jerusalem.

Havnes, Tarjei, and Magne Mogstad. 2011. "No Child Left Behind: Subsidized Child Care and Children's Long-Run Outcomes." *American Economic Journal: Economic Policy* 3 (2): 97–129.

Havnes, Tarjei, and Magne Mogstad. 2015. "Is universal child care leveling the playing field?" *Journal of Public Economics* 127: 100–114.

Heckman, James, Rodrigo Pinto, and Peter Savelyev. 2013. "Understanding the Mechanisms Through Which an Influential Early Childhood Program Boosted Adult Outcomes." *American Economic Review* 103 (6): 2052–2086.

Hendren, Nathaniel, and Ben Sprung-Keyser. 2020. "A unified welfare analysis of government policies." *Quarterly Journal of Economics* 135 (3): 1209–1318.

Humphries, John Eric, Christopher Neilson, Xiaoyang Ye, and Seth Zimmerman. 2024. "Parents' Earnings and the Returns to Universal Pre-Kindergarten." w33038, National Bureau of Economic Research. Cambridge, MA.

Imbens, Guido W., and Joshua D. Angrist. 1994. "Identification and Estimation of Local Average Treatment Effects." *Econometrica* 62 (2): 467–475.

Jackson, C Kirabo. 2018. "What Do Test Scores Miss? The Importance of Teacher Effects on Non–Test Score Outcomes." *Journal of Political Economy* 126 (5): 2072–2107.

Jackson, C. Kirabo, Julia A Turner, and Jacob Bastian. 2025. "Universal Pre-K as Economic Stimulus: Evidence from Nine States and Large Cities in the U.S.." Working Paper 33767, National Bureau of Economic Research.

Jayachandran, Seema, and Rohini Pande. 2017. "Why Are Indian Children So Short? The Role of Birth Order and Son Preference." *American Economic Review* 107 (9): 2600–2629.

Johnson, Rucker C., and C. Kirabo Jackson. 2019. "Reducing Inequality through Dynamic Complementarity: Evidence from Head Start and Public School Spending." *American Economic Journal: Economic Policy* 11 (4): 310–349.

Karbownik, Krzysztof, and Anthony Wray. 2025. "Lifetime and Intergenerational Consequences of Poor Childhood Health." *Journal of Human Resources* 60 (1): 187–223.

Kimhi, Ayal. 2012. "Pre-primary Education in Israel: Organizational and Demographic Aspects." Policy Report, Taub Center for Social Policy Studies in Israel. Jerusalem.

Kline, Patrick, and Christopher R. Walters. 2016. "Evaluating Public Programs with Close Substitutes: The Case of HeadStart." *The Quarterly Journal of Economics* 131 (4): 1795–1848.

Knudsen, Eric I., James J. Heckman, Judy L. Cameron, and Jack P. Shonkoff. 2006. "Economic, Neurobiological, and Behavioral Perspectives on Building America's Future Workforce." *Proceedings of the National Academy of Sciences* 103 (27): 10155–10162.

Kottelenberg, Michael J., and Steven F. Lehrer. 2017. "Targeted or Universal Coverage? Assessing Heterogeneity in the Effects of Universal Child Care." *Journal of Labor Economics* 35 (3): 609–653.

Lavy, Victor, and Edith Sand. 2018. "On the Origins of Gender Gaps in Human Capital: Short- and Long-term Consequences of Teachers' Biases." *Journal of Public Economics* 167: 263–279.

Lipsey, Mark W., Dale C. Farran, and Kelley Durkin. 2018. "Effects of the Tennessee Prekindergarten Program on Children's Achievement and Behavior Through Third Grade." *Early Childhood Research Quarterly* 45: 155–176.

Lleras-Muney, Adriana, Joseph Price, and Dahai Yue. 2022. "The association between educational attainment and longevity using individual-level data from the 1940 census." *Journal of Health Economics* 84: 102649.

Ludwig, J., and D. L. Miller. 2007. "Does Head Start Improve Children's Life Chances? Evidence from a Regression Discontinuity Design." *The Quarterly Journal of Economics* 122 (1): 159–208.

Magnuson, Katherine A., Robert Kelchen, Greg J. Duncan, Holly S. Schindler, Hilary Shager, and Hirokazu Yoshikawa. 2016. "Do the Effects of Early Childhood Education Programs Differ by Gender? A Meta-analysis." *Early Childhood Research Quarterly* 36: 521–536.

Miller, Douglas L., Na'ama Shenhav, and Michel Grosz. 2023. "Selection into Identification in Fixed Effects Models, with Application to Head Start." *Journal of Human Resources* 58 (5): 1523–1566.

Miller, Sarah, Laura Wherry, and Gloria Aldana. 2022. "Covering Undocumented Immigrants: The Effects of a Large-Scale Prenatal Care Intervention." w30299, National Bureau of Economic Research. Cambridge, MA.

Montroy, Janelle J., Ryan P. Bowles, Lori E. Skibbe, Megan M. McClelland, and Frederick J. Morrison. 2016. "The development of self-regulation across early childhood." *Developmental Psychology* 52 (11): 1744–1762.

Oster, Emily. 2019. "Unobservable Selection and Coefficient Stability: Theory and Evidence." *Journal of Business & Economic Statistics* 37 (2): 187–204.

Pages, Remy, Dylan J. Lukes, Drew H. Bailey, and Greg J. Duncan. 2020. "Elusive Longer-Run Impacts of Head Start: Replications Within and Across Cohorts." *Educational Evaluation and Policy Analysis* 42 (4): 471–492.

Pollakowski, Henry O., Daniel H. Weinberg, Fredrik Andersson, John C. Haltiwanger, Giordano Palloni, and Mark J. Kutzbach. 2022. "Childhood Housing and Adult Outcomes: A Between-Siblings Analysis of Housing Vouchers and Public Housing." *American Economic Journal: Economic Policy* 14 (3): 235–272.

Reardon, Sean F., Erin M. Fahle, Demetra Kalogrides, Anne Podolsky, and Rosalía C. Zárate. 2019. "Gender Achievement Gaps in U.S. School Districts." *American Educational Research Journal* 56 (6): 2474–2508.

Reeves, Richard V. 2022. *Of Boys and Men: Why the Modern Male is Struggling, Why it Matters, and What to Do About it*. Washington, D.C: Brookings Institution Press.

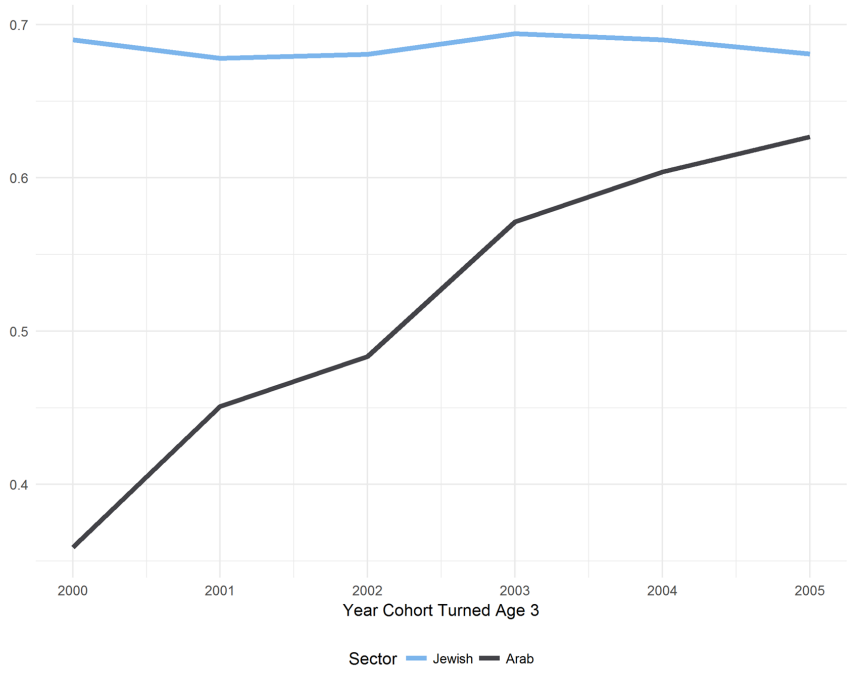
Tandon, Pooja S., Chuan Zhou, and Dimitri A. Christakis. 2012. "Frequency of Parent-Supervised Outdoor Play of US Preschool-Aged Children." *Archives of Pediatrics & Adolescent Medicine* 166 (8).

UNDP. 2006. "The Arab Human Development Report: Towards the Rise of Women in the Arab World." UN Report, UNDP, New York.

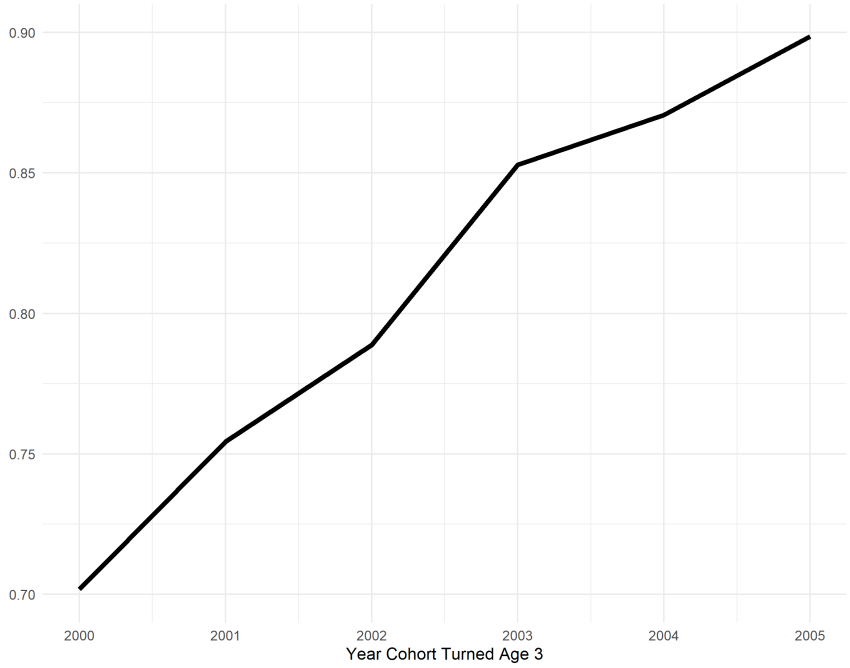
Weiland, Christina, Rebecca Unterman, Anna Shapiro, Sara Staszak, Shana Rochester, and Eleanor Martin. 2020. "The Effects of Enrolling in Oversubscribed Prekindergarten Programs Through Third Grade." *Child Development* 91 (5): 1401–1422.

Yassine-Hamdan, Nahla, and John Strate. 2020. "Gender Inequality in the Arab World: A Comparative Perspective." *Contemporary Arab Affairs* 13 (3): 25–50.

Figures



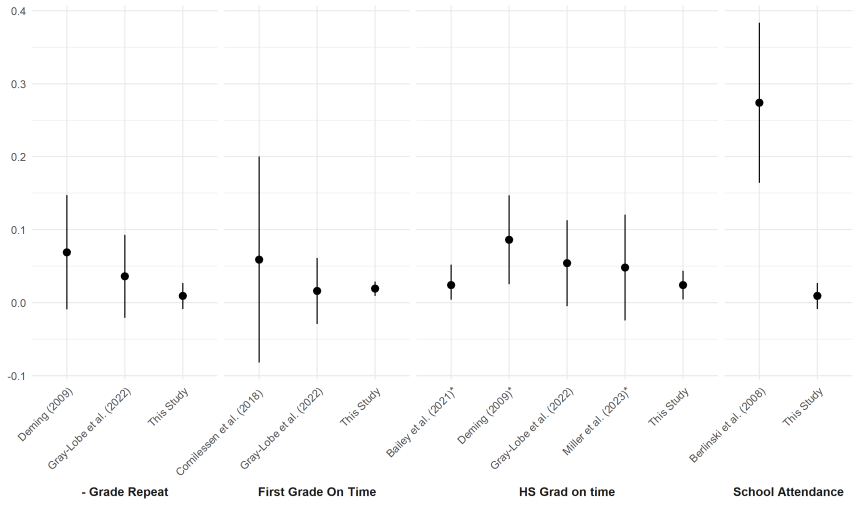
(a) Enrollment Rate in Pre-K at Age 3 by Year and Sector



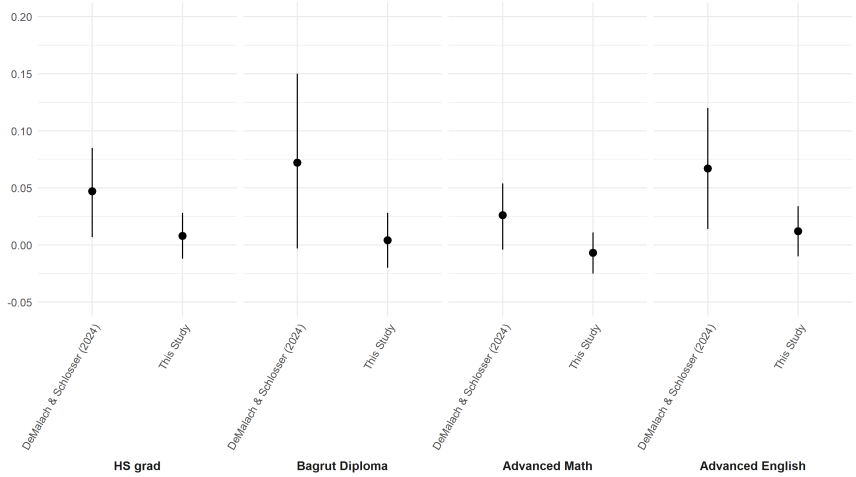
(b) Enrollment Rate in Pre-K at Age 3 in Sample Municipalities

Figure 1: Enrollment Rate in Pre-K at Age 3

Notes. The figure depicts enrollment rates in public pre-K for age three over time. Panel A shows rates for the universe of Arab (black line) and Jewish students. Panel B shows enrollment rate for the Arab localities included in the 2001 special order.



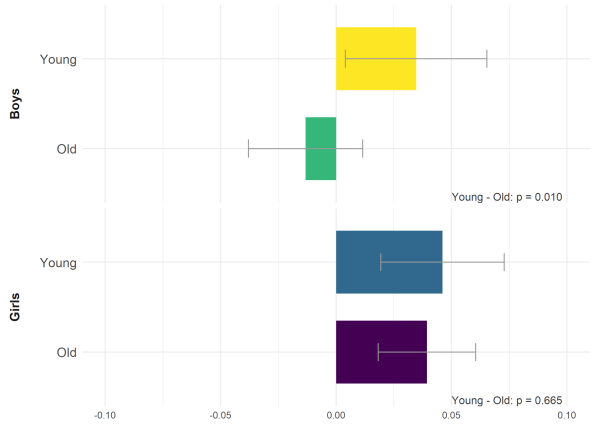
Panel A: Comparison to Other Pre-K Treatment Effect Estimates



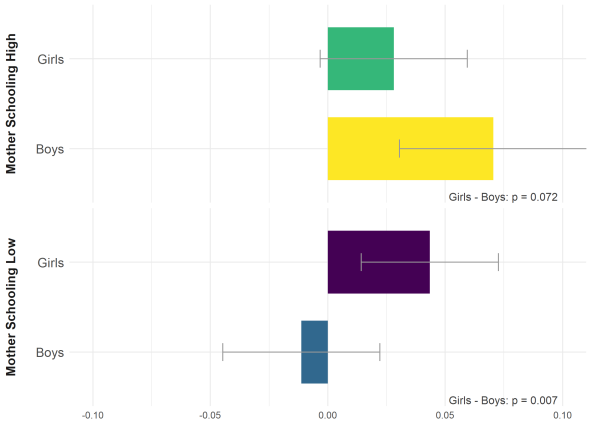
Panel B: Comparison to Estimates in DeMalach and Schlosser (2026)

Figure 2: Effects of Large-Scale Preschool Programs Across Studies

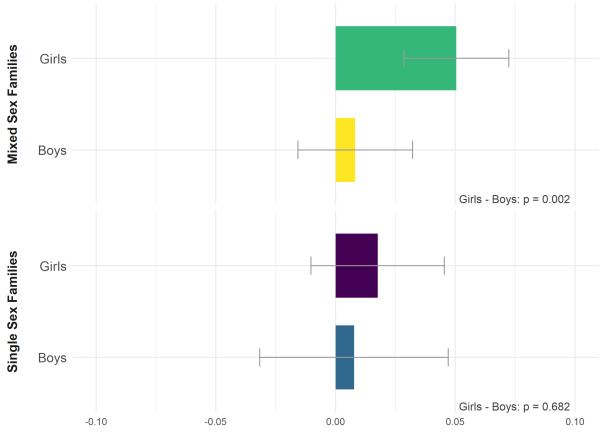
Notes. Panel A plots point estimates and 95% confidence intervals for treatment effects on four outcomes from different studies. Effects on grade repetition are multiplied by negative one for ease of presentation. For on-time high school graduation, an asterisk indicates that an on-time estimate was not available and an ever-graduated-high-school outcome is used instead. For school attendance, I use this study's estimate of ever attending 11th grade, and for Berlinski, Galiani, and Manacorda (2008), the estimate of attending school at age 15. For Miller, Shenhav, and Grosz (2023) I use the estimate with Head Start participants as the target group. Panel B shows treatment effect estimates and 95% confidence intervals from this study and DeMalach and Schlosser (2026). To calculate treatment point estimates, I scale the ITT estimate by the first-stage estimate reported in DeMalach and Schlosser (2026). To construct confidence intervals, I first calculate the implied 95% confidence interval for the ITT from the reported ITT standard error and then scale it by the first stage.



Panel A: Sex and Relative-Age Treatment Heterogeneity



Panel B: Sex and Maternal-Schooling Treatment Heterogeneity



Panel C: Sex and Sibling-Sex-Composition Treatment Heterogeneity

Figure 3: Estimates of Heterogeneous Effect on Graduating from High School on Time

Notes. The figure plots estimates of β in equation 1 for different subgroups. In each panel, estimates are obtained by interacting pre-K with a full set of dummy variables for each subgroup. In Panel A, "young" refers to children born September–December and "old" to those born January–August, following the school-entry-age cutoff. Each panel reports p-values from tests of equality of the effects shown: girls versus boys within each subgroup in Panels B and C, and young versus old children of each sex in Panel A.

Tables

Table 1: Summary Statistics

	Entire Sample (1)	Family FE Sample (2)
Attended Pre-K	0.815	0.806
	(0.388)	(0.395)
Girl	0.489	0.499
	(0.500)	(0.500)
Father Schooling	10.296	10.241
	(2.778)	(2.723)
Mother Schooling	10.092	10.008
	(2.611)	(2.555)
Family Size	3.784	3.915
	(1.954)	(1.973)
Month of Birth	6.621	6.626
	(3.418)	(3.423)
Birth Order	3.144	3.161
	(2.001)	(1.995)
First Grade on Time	0.980	0.979
	(0.141)	(0.143)
Graduated High School on Time	0.826	0.820
	(0.379)	(0.385)
Graduated High School	0.858	0.852
	(0.349)	(0.355)
Attended 11th Grade	0.884	0.879
	(0.321)	(0.326)
Ever Repeated a Grade	0.106	0.108
	(0.308)	(0.310)
Diploma (Bagrut)	0.540	0.527
	(0.498)	(0.499)
Diploma Credits	19.745	19.403
	(12.840)	(12.872)
Advanced Math	0.223	0.213
	(0.417)	(0.409)
Advanced English	0.483	0.466
	(0.500)	(0.499)
Advanced Hebrew	0.316	0.312
	(0.465)	(0.463)
Number of Students	54,264	40,644
Number of Families	30,466	16,747

Notes. The table shows summary statistics for key background and outcome variables. Column 1 shows statistics for students in birth cohorts 1996 to 2001 who resided in sample localities. Column 2 shows results for children who have a sibling in the sample.

Table 2: Predictors of Public Pre-K Enrollment

	Dependent variable: Pre-K Enrollment			
	Univariate	Multivariate Models		
	(1)	(2)	(3)	(4)
Girl	−0.001	0.001	−0.001	0.004
	(0.003)	(0.003)	(0.003)	(0.004)
Father Schooling > 10	0.014***	0.009***	0.019***	-
	(0.003)	(0.003)	(0.004)	
Mother Schooling > 10	0.118***	0.068***	0.029***	-
	(0.003)	(0.004)	(0.008)	
Family Size	−0.033***	−0.039***	−0.003	-
	(0.001)	(0.003)	(0.003)	
Month of Birth	−0.007***	−0.007***	−0.007***	−0.007***
	(0.0005)	(0.0005)	(0.001)	(0.001)
Birth Order	−0.025***	0.024***	−0.004	0.003
	(0.001)	(0.002)	(0.003)	(0.003)
Municipality–Cohort Fixed Effects	✗	✗	✓	✓
Family Fixed Effects	✗	✗	✗	✓
Observations	54,264	54,264	54,264	40,644

Notes. The table presents OLS estimates from regressions of public pre-K enrollment on various characteristics. Column 1 reports coefficients from separate univariate regressions in which each characteristic is the sole independent variable. Columns 2–4 report multivariate specifications: without municipality-cohort fixed effects (column 2), with municipality-cohort fixed effects (column 3), and with municipality-cohort and family fixed effects (column 4). The sample includes all children residing in sample municipalities from birth cohorts 1996–2001. Robust standard errors are reported in parentheses below each coefficient; in column 4, standard errors are clustered at the family level. *p<0.1; **p<0.05; ***p<0.01.

Table 3: Characteristics of Switchers and Non-switchers

	Non-switchers		Switchers	
	Never Takers	Always Takers	Expansion Compliers	Expansion Defiers
	(1)	(2)	(3)	(4)
Number of Families	1,430	11,982	2,608	727
Children in Sample per Family	2.58	2.37	2.63	2.47
Father Years of Schooling	9.87	10.44	9.92	10.08
Mother Years of Schooling	9.09	10.37	9.53	9.61
Young Child Born Sep-Dec (%)	44.8%	38.8%	35.3%	50.6%
Birth Order	3.99	2.83	3.69	3.50
Family Size	4.82	3.49	4.45	4.19
Oldest Child Attended				
Pre-K at Age 3	0	1	0	1
Pre-K at Age 4	0.40	0.97	0.73	0.85

Notes. The table reports means of selected family characteristics by four mutually exclusive groups of families based on pre-K enrollment status of siblings who are included in the sample: neither the oldest nor youngest sampled sibling attended (column 1), both the oldest and youngest sampled siblings attended (column 2), the youngest sampled sibling attended but the oldest did not (column 3), and the oldest sampled sibling attended but the youngest did not (column 4).

Table 4: Effects of Public Pre-K on School Progression and High School Achievements

	Panel A: School-Progression Outcomes				
	First Grade on time	Grade Repeat	11th Grade	Grad HS on Time	Grad HS
	(1)	(2)	(3)	(4)	(5)
Estimate	0.019***	−0.013*	0.009	0.024***	0.008
	(0.004)	(0.007)	(0.007)	(0.008)	(0.007)
Mean	0.979	0.108	0.879	0.820	0.852
Observations	40,644	40,644	40,644	40,644	40,644
Number of Families	16,747	16,747	16,747	16,747	16,747
Adjusted R^2	0.077	0.237	0.293	0.305	0.309
	Panel B: HS Diploma Outcomes				
	HS Diploma (Bagrut)	Diploma Credits	Advanced Math	Advanced English	Advanced Hebrew
	(1)	(2)	(3)	(4)	(5)
Estimate	0.003	0.442**	−0.007	0.012	0.010
	(0.009)	(0.215)	(0.007)	(0.009)	(0.008)
Mean	0.527	19.403	0.213	0.466	0.312
Observations	40,644	40,644	40,644	40,644	40,644
Number of Families	16,747	16,747	16,747	16,747	16,747
Adjusted R^2	0.396	0.549	0.436	0.476	0.439

Notes. The table shows estimates of β in equation 1. Panel A shows estimates of the effects on school-progression outcomes. Panel B shows estimates of the effects on educational outcomes measured at the end of high school. All models control for sex, month of birth, birth order, and municipality-specific cohort fixed effects. Standard errors, clustered at the family level, are displayed in parentheses. *p<0.1; **p<0.05; ***p<0.01.

Table 5: Robustness Tests

Robustness Test	**HS Grad On Time**	**HS Diploma (Bagrut)**
No Expansion Defiers	0.020**	−0.005
	(0.009)	(0.011)
Number of Observations	38,846	
Common Cohort Fixed Effect	0.025***	−0.005
	(0.007)	(0.009)
Number of Observations	40,644	
No Control for Month of Birth	0.028***	0.006
	(0.008)	(0.009)
Number of Observations	40,644	
All Arab-majority Municipalities	0.015***	0.001
	(0.004)	(0.005)
Number of Observations	133,751	
Only Youngest and Oldest Child	0.023**	0.005
	(0.009)	(0.011)
Number of Observations	34,182	

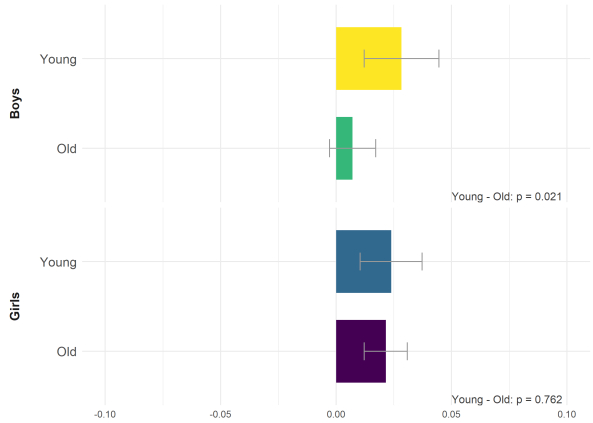
Notes. The table shows estimates of β for high school graduation on time (column 1) and obtaining a high school diploma (Bagrut) (column 2) across five robustness tests: excluding expansion defiers from the sample; replacing municipality-specific cohort fixed effects with common cohort fixed effects; removing controls for month of birth; including all Arab-majority municipalities; and keeping only the oldest and youngest children in the sample. Standard errors, clustered at the family level, are reported in parentheses. *p<0.1; **p<0.05; ***p<0.01.

Table 6: Effects of Public Pre-K on School Progression for Subgroups

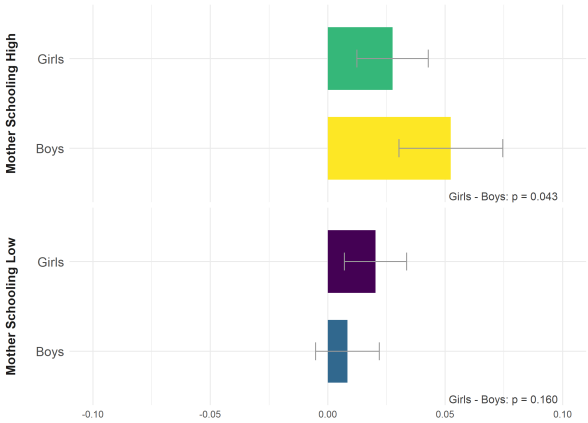
	By Sex		By Relative Age		By Mother Education		By Family Size	
	Boys (1)	Girls (2)	Young (3)	Old (4)	High (5)	Low (6)	Small (7)	Large (8)
First Grade on Time	0.016***	0.022***	0.026***	0.014***	0.039***	0.014**	0.016***	0.021***
	(0.005)	(0.004)	(0.006)	(0.004)	(0.007)	(0.005)	(0.005)	(0.005)
p-value for equal effects	0.240		0.043		0.003		0.401	
Grade Repeat	−0.007	−0.018**	−0.018*	−0.009	−0.021*	−0.017	−0.008	−0.016*
	(0.009)	(0.008)	(0.009)	(0.009)	(0.012)	(0.011)	(0.009)	(0.009)
p-value for equal effects	0.240		0.369		0.810		0.472	
11^{th} Grade	−0.017*	0.036***	0.016*	0.005	0.016	0.015	−0.011	0.023***
	(0.009)	(0.008)	(0.009)	(0.008)	(0.011)	(0.011)	(0.009)	(0.009)
p-value for equal effects	0.000		0.263		0.978		0.003	
Grad HS on Time	0.006	0.042***	0.040***	0.013	0.050***	0.016	0.020*	0.027***
	(0.011)	(0.009)	(0.011)	(0.011)	(0.014)	(0.012)	(0.011)	(0.009)
p-value for equal effects	0.002		0.026		0.057		0.595	
Grad HS	−0.014	0.030***	0.011	0.005	0.024**	0.001	0.003	0.011
	(0.010)	(0.009)	(0.010)	(0.009)	(0.012)	(0.011)	(0.010)	(0.009)
p-value for equal effects	0.000		0.628		0.131		0.503	
Joint p-value across outcomes	6.03e−10		0.0448		0.0121		0.0516	

Notes. This table reports estimates of β in equation 1 on school-progression outcomes for subgroups. Columns 1 and 2 compare estimates for boys and girls; columns 3 and 4 display estimates by relative age; columns 5 and 6 show estimates by maternal education (more or less than 10 years); and columns 7 and 8 show estimates by family size (larger than three children versus smaller than or equal to three). p-values for equal effects come from tests of the null hypothesis that effects are equal across subgroups. All models control for sex, month of birth, birth order, and municipality-specific cohort fixed effects. The last row shows joint p-values from tests of the null hypothesis that subgroups have equal effects across all outcomes in the table. *p<0.1; **p<0.05; ***p<0.01.

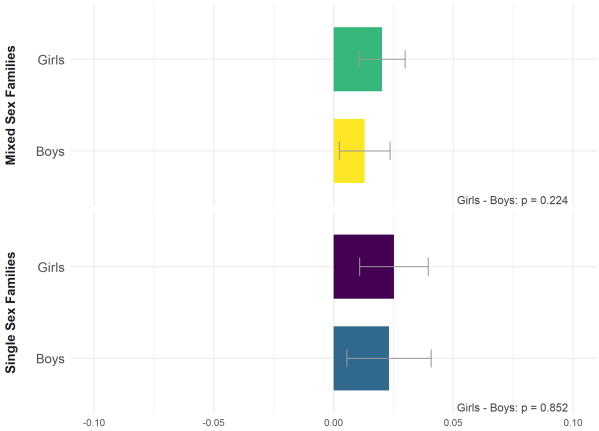
Appendix



Panel A: Sex and Relative Age Treatment Heterogeneity



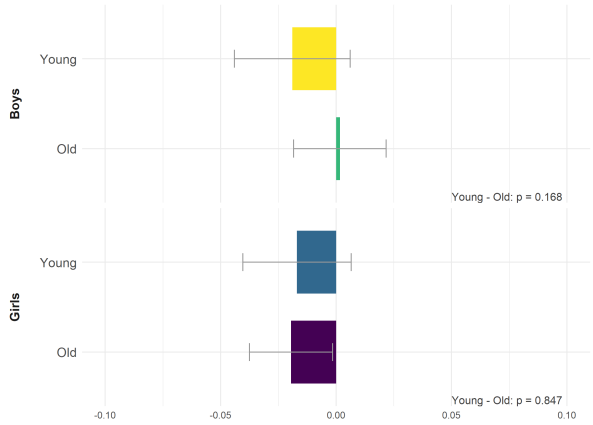
Panel B: Sex and Maternal Schooling Treatment Heterogeneity



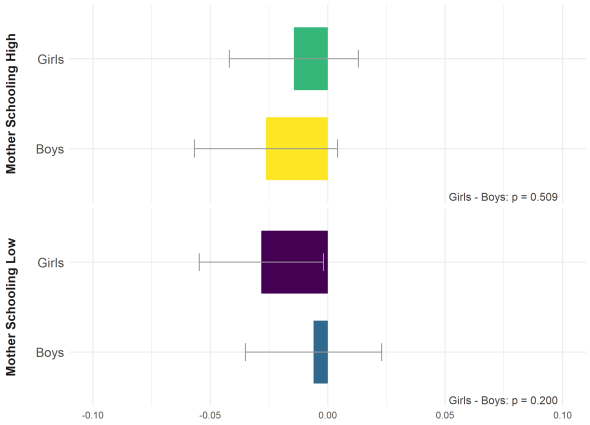
Panel C: Sex and Sibling Sex Composition Treatment Heterogeneity

Figure A1: Estimates of Heterogeneous Effect on First Grade on Time

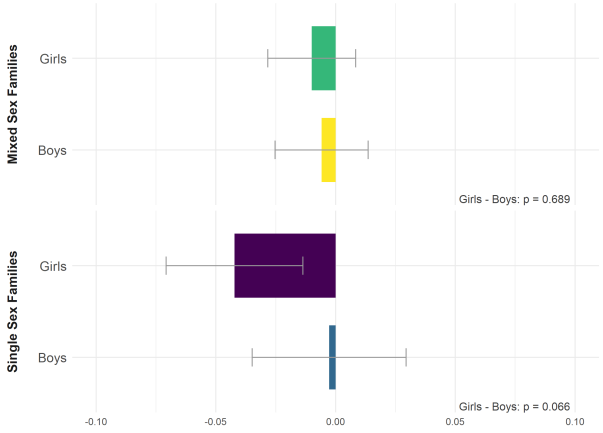
Notes. The figure plots estimates of β in equation 1 for different subgroups. In each panel estimates are obtained by interacting pre-K with a full set of dummy variables for each subgroup. Each panel reports p-values from tests of equality of the effects shown: girls versus boys within each subgroup in Panels B and C, and young versus old children of each sex in Panel A.



Panel A: Sex and Relative Age Treatment Heterogeneity



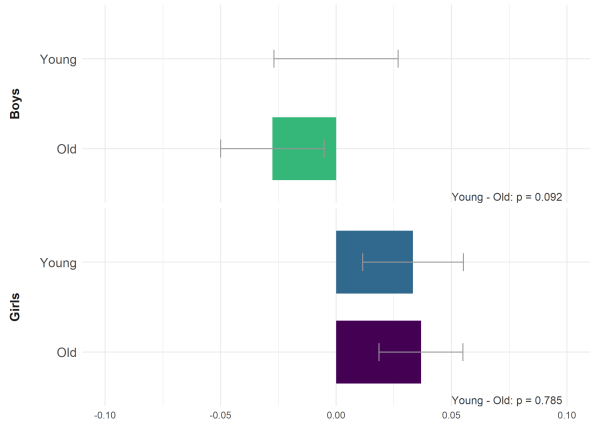
Panel B: Sex and Maternal Schooling Treatment Heterogeneity



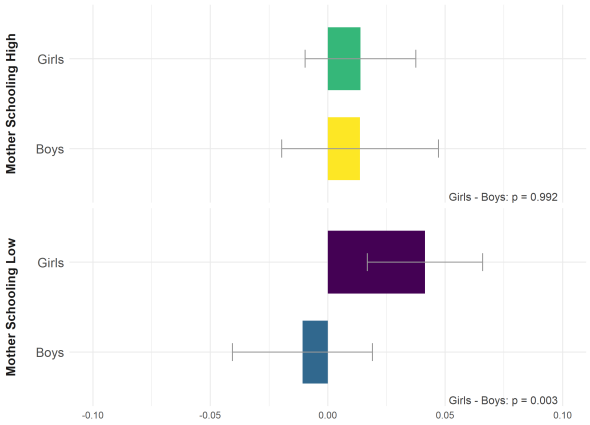
Panel C: Sex and Sibling Sex Composition Treatment Heterogeneity

Figure A2: Estimates of Heterogeneous Effect on Ever Repeating a Grade

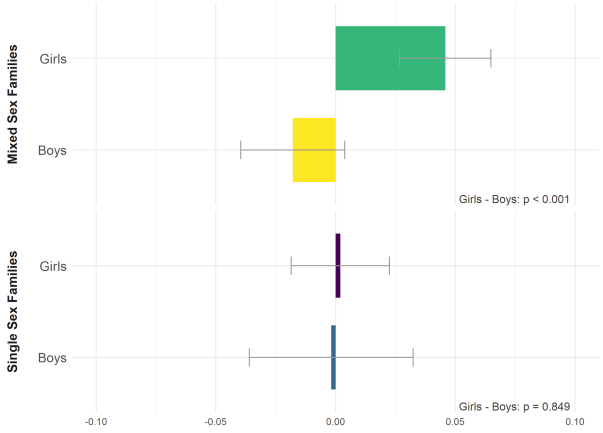
Notes. The figure plots estimates of β in equation 1 for different subgroups. In each panel estimates are obtained by interacting pre-K with a full set of dummy variables for each subgroup. Each panel reports p-values from tests of equality of the effects shown: girls versus boys within each subgroup in Panels B and C, and young versus old children of each sex in Panel A.



Panel A: Sex and Relative Age Treatment Heterogeneity



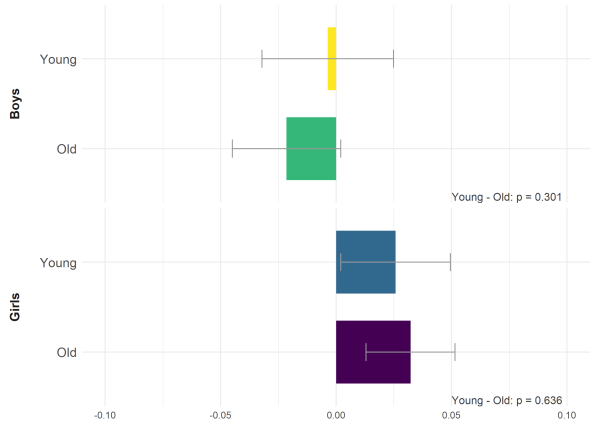
Panel B: Sex and Maternal Schooling Treatment Heterogeneity



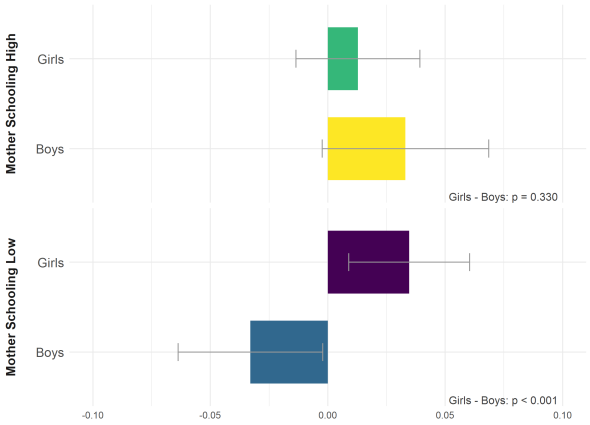
Panel C: Sex and Sibling Sex Composition Treatment Heterogeneity

Figure A3: Estimates of Heterogeneous Effect on Attending 11th Grade

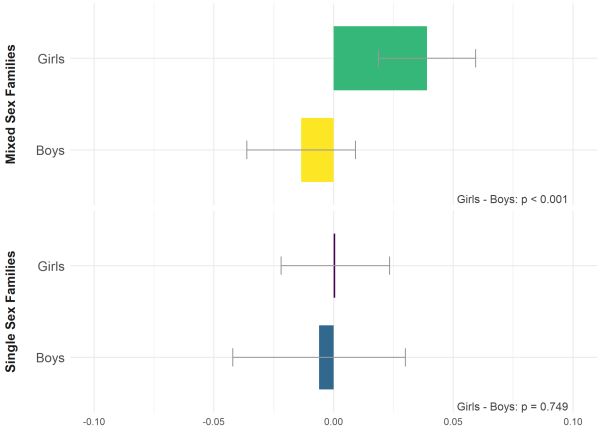
Notes. The figure plots estimates of β in equation 1 for different subgroups. In each panel estimates are obtained by interacting pre-K with a full set of dummy variables for each subgroup. Each panel reports p-values from tests of equality of the effects shown: girls versus boys within each subgroup in Panels B and C, and young versus old children of each sex in Panel A.



Panel A: Sex and Relative Age Treatment Heterogeneity



Panel B: Sex and Maternal Schooling Treatment Heterogeneity



Panel C: Sex and Sibling Sex Composition Treatment Heterogeneity

Figure A4: Estimates of Heterogeneous Effect on Ever Graduating from High School

Notes. The figure plots estimates of β in equation 1 for different subgroups. In each panel estimates are obtained by interacting pre-K with a full set of dummy variables for each subgroup. Each panel reports p-values from tests of equality of the effects shown: girls versus boys within each subgroup in Panels B and C, and young versus old children of each sex in Panel A.

Table A1: Estimates of the Impact of Public Pre-K Attendance on Obtaining a High School Diploma (*Bagrut*)

	Dependent Variable: High School Diploma (*Bagrut*)				
	(1)	(2)	(3)	(4)	(5)
PreK-3	0.124***	0.089***	0.056***	0.050***	0.003
	(0.006)	(0.007)	(0.006)	(0.007)	(0.009)
Background Controls	✗	✓	✓	✓	✓
Municipality-Cohort FE	✗	✗	✓	✓	✓
Siblings Sample	✗	✗	✗	✓	✓
Family FE	✗	✗	✗	✗	✓
Observations	54,264	54,264	54,264	40,644	40,644
Adjusted R^2	0.009	0.058	0.195	0.194	0.396

Notes. This table reports coefficients from regressing high school diploma on public pre-K attendance at age three. Column 1 presents the estimated coefficient from a univariate model. Column 2 adds controls for sex, month of birth, birth order, and parental education. Columns 3 and 4 also incorporate municipality-cohort fixed effects. Column 5 further adds family fixed effects. The sample in columns 1–3 includes all children residing in sample municipalities from birth cohorts 1996-2001. Columns 4–5 restrict the sample to children with at least one sibling in the sample. Robust standard errors are reported in parentheses below each coefficient. *p<0.1; **p<0.05; ***p<0.01.

Table A2: Robustness of Pre-K Effect: Reweighted ATE Estimates

	Panel A: School-Progression Outcomes				
	First Grade on time	Grade Repeat	11th Grade	Grad HS on Time	Grad HS
	(1)	(2)	(3)	(4)	(5)
Estimate	0.025***	−0.018*	0.013	0.033***	0.017
	(0.007)	(0.010)	(0.010)	(0.012)	(0.010)
	Panel B: HS Diploma Outcomes				
	HS Diploma (Bagrut)	Diploma Credits	Advanced Math	Advanced English	Advanced Hebrew
	(1)	(2)	(3)	(4)	(5)
Estimate	−0.001	0.724***	−0.010	0.015	0.019
	(0.014)	(0.334)	(0.012)	(0.014)	(0.013)

Notes. The table reports estimates of β from Equation 1, using observation weights designed to recover the average treatment effect following the reweighting method proposed by Miller, Shenhav, and Grosz (2023). The method was implemented in two stages. First, a multinomial logit model was estimated to predict the probability that each family belongs to one of four groups defined by the interaction of switching status and pre-K attendance. In the second stage, weights were computed as the ratio of the predicted probability of being in the attendance group to the probability of being in the switcher group. This ratio was multiplied by the inverse of the family's conditional variance of pre-K attendance. These weights were then used to estimate Equation 1. For additional details, see the notes to Table 4. *p<0.1; **p<0.05; ***p<0.01.

Table A3: Effects of Public Pre-K on High School Achievements for Subgroups

	By Sex		By Relative Age		By Mother Education		By Family Size	
	Boys (1)	Girls (2)	Young (3)	Old (4)	High (5)	Low (6)	Small (7)	Large (8)
HS Diploma (*Bagrut*)	0.003	0.002	0.009	-0.001	-0.013	-0.009	0.002	0.003
	(0.011)	(0.011)	(0.012)	(0.010)	(0.018)	(0.014)	(0.013)	(0.011)
p-value for equal effects	0.916		0.472		0.871		0.940	
Total Units	0.277	0.608**	0.499*	0.405	0.764*	0.307	0.432	0.446*
	(0.272)	(0.261)	(0.291)	(0.247)	(0.411)	(0.316)	(0.301)	(0.262)
p-value for equal effects	0.296		0.766		0.353		0.969	
Advanced Math	-0.013*	-0.000	-0.005	-0.008	-0.013	-0.005	-0.003	-0.010
	(0.008)	(0.009)	(0.009)	(0.008)	(0.016)	(0.010)	(0.010)	(0.008)
p-value for equal effects	0.200		0.772		0.637		0.583	
Advanced English	0.010	0.014	0.016	0.010	0.020	0.013	0.014	0.011
	(0.010)	(0.011)	(0.012)	(0.010)	(0.017)	(0.013)	(0.013)	(0.010)
p-value for equal effects	0.741		0.627		0.752		0.879	
Advanced Hebrew	0.019*	0.001	0.023**	0.001	-0.020	0.029**	0.010	0.010
	(0.010)	(0.011)	(0.011)	(0.010)	(0.018)	(0.012)	(0.012)	(0.010)
p-value for equal effects	0.140		0.079		0.019		0.973	
Joint p-value across outcomes	0.412		0.547		0.241		0.997	

Notes. This table reports estimates of β in equation 1 on high school educational achievements for subgroups. Columns (1) and (2) compare estimates for boys and girls; columns (3) and (4) display estimates by relative age; columns (5) and (6) show estimates by maternal education (more or less than 10 years); and columns (7) and (8) show estimates by family size (larger than three versus smaller than or equal to three). All models control for sex, month of birth, birth order, and municipality-specific cohort fixed effects. *p<0.1; **p<0.05; ***p<0.01.

Table A4: Association Between Pre-K Enrollment and Child and Maternal Characteristics

	Panel A: Child Birth Weight		
	Mean [SD]	No Family FE	Family FE
	(1)	(2)	(3)
Birth weight (kg)	3.23	0.012***	0.004
	[0.52]	(0.002)	(0.003)
Family Fixed Effects		✗	✓
Number of children		144,348	
Number of families		61,913	
	Panel B: Maternal Labor Supply		
	Mean [SD]	No Family FE	Family FE
	(1)	(2)	(3)
Mother worked	0.27	0.060***	0.003
	[0.44]	(0.003)	(0.004)
Months worked	2.39	0.0054***	0.0004
	[4.41]	(0.0003)	(0.0004)
Mother's earnings	10.11	0.0008***	−0.00007
(in NIS 1,000)	[21.45]	(0.0001)	(0.0001)
Family Fixed Effects		✗	✓
Number of children		151,773	
Number of families		66,441	

Notes. The table examines the relationship between pre-K enrollment at age three and child birth weight (Panel A) and maternal labor market outcomes when the child was three years old (Panel B). The sample includes Arab children born 2002–2007 with at least one sibling born in the same period. Panel A restricts to children with non-missing birth weight. Column (1) reports means with standard deviations in brackets. Columns (2) and (3) report coefficients from regressions of each outcome on a pre-K attendance indicator. All specifications include municipality-by-cohort fixed effects, birth-order fixed effects, and a female indicator. Column (3) adds family fixed effects. Standard errors are clustered at the family level. Data are from the Israeli Central Bureau of Statistics and cover a different period than the main analysis (birth cohorts 1996–2001). *p<0.1; **p<0.05; ***p<0.01.

Table A5: Sibling Spillovers from Pre-K Attendance: Effect on On-Time High School Graduation

	Panel A: Heterogeneity by Sibling Spacing	
	(1)	(2)
PreK	0.023	0.024
	(0.008)	(0.015)
PreK × Age Gap	0.007	–
	(0.009)	
PreK ×1{Age Gap > Median}	–	-0.005
		(0.019)
Number of students	34,182	34,182
Number of families	16,747	16,747
	Panel B: Difference-in-Differences	
	Main Sample	All Arab-majority Towns
	(1)	(2)
Child 2 × Treated	0.001	-0.007
	(0.020)	(0.008)
Number of students	8,966	29,916
Number of families	4,483	14,966

Notes. This table examines whether one sibling's pre-K attendance affects another's likelihood of graduating high school on time. Panel A reports the estimate of β and the coefficient on the interaction between pre-K attendance and the age gap between the youngest and oldest sibling in each family (column 1), and the interaction with a dummy indicating whether the age gap exceeds the sample median (2.11 years) (column 2). To facilitate interpretation of the coefficient in column (1), the age gap variable is standardized using the sample mean and standard deviation of the age gap. The sample for these regressions is restricted to the oldest and youngest child from each family. Panel B reports estimates of θ from equation 2. The sample includes first- and second-born children in families with three children in the sample. *Treated* is an indicator for whether the third-born child attended pre-K. Column (1) uses the main analysis sample, while column (2) uses the full population of children residing in Arab-majority municipalities. In both panels, the dependent variable is an indicator for on-time high school graduation, and standard errors in parentheses are clustered at the family level.