Experimental Evidence on the Effect of Full-Day Pre-Kindergarten on Early Elementary Special Education Designations

Danielle Shaw Attaway, Katharine Parham Malhotra & Allison Atteberry

View supplementary material

Published online: 19 Aug 2025.

Submit your article to this journal

Article views: 2

View related articles

View Crossmark data

This article has been awarded the Centre for Open Science 'Preregistered' badge.

Routledge
Taylor & Francis Group

INTERVENTION, EVALUATION, AND POLICY STUDIES

Check for updates

Experimental Evidence on the Effect of Full-Day Pre-Kindergarten on Early Elementary Special Education Designations

Danielle Shaw Attaway[a], Katharine Parham Malhotra[b], and Allison Atteberry[b]

[a]American Institutes for Research, Arlington, VA, USA; [b]School of Education and Human Development, University of Virginia, Charlottesville, VA, USA

ABSTRACT

School districts face a difficult decision when considering a move from half- to full-day pre-K, partly because this is a relatively expensive policy option—each half-day classroom/teacher can serve twice as many students (AM/PM) than one full-day classroom. Yet full-day pre-K could confer meaningful benefits to students, by increasing children's exposure to foundational experiences that support their early development and increasing the time children spend with their pre-K teachers. We explore whether full-day pre-K reduces later special education (SE) designations in ways that could represent a cost-savings to districts, without reducing access to SE services that would have benefited students. Leveraging a randomized field experiment, we estimate the causal effect of full- versus half-day pre-K and find that full-day pre-K leads to a nearly 50% reduction in the likelihood of having an SE designation in early elementary grades among children who, at baseline, exhibit a possibility of benefiting from developmental supports. We also find important differential effects of full- versus half-day pre-K between children whose primary home language is or is not English, as well as male and female children.

Introduction

In this article, we examine whether randomly assigned offers of full- versus half-day pre-kindergarten (pre-K) affect students' likelihood of having special education (SE) designations in early elementary grades Kindergarten (K) through 3. Full- versus half-day pre-K represents a meaningful difference in pre-K program intensity. As such, it has the potential to shape children's developmental trajectories by increasing their exposure to both instructional and non-instructional experiences that support early learning and development (Duncan & Magnuson, 2013; Halfon & Hochstein, 2002). The greater time spent in pre-K may allow teachers to more deeply understand the children in their classroom academically, socially, and developmentally and use this knowledge to influence children's future academic paths (Chetty et al., 2014; Jackson, 2018),

including potentially altering odds of receiving an SE designation in early elementary grades.

Of course, the majority of children are never candidates for SE designations; for these children, we should expect to find no effect of full- vs. half-day pre-K on SE designation rates. We therefore seek to confirm that any apparent effects on SE designations follow this logical pattern—that is, concentrated among students who are assessed at baseline as having any possible need for developmental supports. Where possible, we also explore whether effects vary by key student characteristics sometimes associated with over-representation[1] in SE—such as being an English language learner or male student—to better understand the pathways through which full-day pre-K may influence later SE identification.

Atteberry et al. (2019) conducted a randomized control trial (RCT) in Colorado's Westminster Public School District (WPS) that was designed to examine the causal effects of program intensity (full-vs. half-day pre-K) on a variety of child and family outcomes. An earlier published study that focused on children's school readiness outcomes finds that full-day pre-K offers had substantial and positive short-term effects, relative to half-day offers, on a variety of end-of-pre-K child assessments including the Peabody Picture Vocabulary Test (PPVT), the Early Screening Inventory-Revised (ESI-R), and the Teaching Strategies GOLD (TS-GOLD) assessment—a teacher-reported assessment of children's cognitive, literacy, math, and physical development (Atteberry et al., 2019). These findings suggest that WPS' new full-day pre-K offerings positively influenced children's early learning by the end of the pre-K year.

From the outset, WPS also hypothesized that full-day pre-K could reduce SE designations in early grades.[2] In fact, its Full-Day Pre-K (FDPK) pilot program was funded through a Pay-for-Success (PFS) model, where "success" repayments were tied to reducing SE designations. WPS envisioned two pathways for how extending pre-K from half- to full-day might achieve this. First, longer exposure to high-quality pre-K could provide more intensive developmental support, reducing the need for common early childhood SE designations (Bai et al., 2020; Magnuson et al., 2016; McCoy et al., 2017). Second, additional time could help teachers develop more nuanced understandings of students' developmental trajectories, thus avoiding common *mis*identifications of students as having special needs, which are more common among young boys and children whose first language is not English (Coutinho & Oswald, 2005; Linn & Hemmer, 2011; Samson & Lesaux, 2009; Sullivan, 2011; Sullivan & Bal, 2013). Alternatively, WPS recognized that full-day pre-K could beneficially increase SE designations if extended observation helped better identify children needing services. Regardless of the direction of the effect, the financial implications of this question are critical. Reducing unnecessary SE designations could free district funds for broader student support, while increasing appropriate designations would ensure resources reach children who need them most.

---

[1]Other student groups, notably Black children and those from low-income backgrounds, have been found to be overrepresented in SE (Morgan, 2020; Schifter et al., 2019). However, our sample includes few Black students and is almost entirely low-income, limiting our ability to estimate effects for these subgroups precisely.

[2]The research team identified special needs diagnosis as an outcome when the study was pre-registered with the Open Science Framework when applying for initial research funding (November 2018), and again with SREE when we applied for additional research funding (Registry ID: 15380.1v1).

The current study leverages this randomized field experiment to estimate the causal effect of full- versus half-day pre-K on the likelihood of having an SE designation in grades K, 1, 2 and 3. We focus on presenting effects of the district's (randomized) *offers* of full- vs. half-day spots—i.e., an intent-to-treat (ITT) estimand. In Online Appendix C, we also introduce and use an instrumental variable approach to estimate effects of actual *uptake* of full-day pre-K—i.e., a treatment-on-treated (TOT) estimand. The pattern of results for TOT effects is quite similar to the ITT findings except 20 to 30% larger, as they are concentrated among compliers; we therefore focus on the ITT effects in this article. In primary analyses, we compare effect estimates between two groups of children based on their baseline scores on the ESI-R, a special needs screening assessment administered by the research team. The upper 50% of the ESI-R distribution includes children who are unlikely to ever need SE services and, therefore, are not expected to exhibit a treatment response. In contrast, children in the lower half have at least some potential need for further screening, making them the group for whom an impact on SE designations might arise.

As a preview of findings, this is indeed what we observe. Among children scoring in the upper half of the ESI-R distribution at baseline, full-day pre-K offers had no effect on subsequent SE designations. For those in the lower half, full-day pre-K offers reduced the likelihood of having a subsequent SE designation in grades K – 2 by between 7.5 and 11.2 percentage points—a sizeable reduction, relative to a base rate of around 20 percent. To explore potential mechanisms, we also compare effects among key student subgroups for whom there is an empirically-based theory for differential effects—based on primary home language, as well as male versus female children. We find, for instance, that SE designation reductions are concentrated among children whose primary home language is not English, for whom full-day offers caused a 10.2 percentage-point reduction in likelihood of having a grade 1 SE designation; no apparent effects were found for children whose primary home language *was* English.

This article proceeds as follows. We first provide the relevant context for this study, background on the underlying experimental design, and this article's contribution to existing work on the impacts of pre-K on SE identification. Given that it is unclear whether full-day pre-K would cause SE designations to increase or decrease in early elementary grades, and whether those changes would be desirable or not, we present a conceptual framework illustrating potential theories of change wherein *either* increases or decreases in SE designations could be beneficial for students. We describe the data used in our analysis as well as details about our specific analytic approach before presenting our main results and subgroup analyses. We conclude by revisiting the conceptual framework and discussing implications for future practice and research.

Contribution and Research Questions

While existing literature suggests that participating in a pre-K program can reduce students' SE placements in elementary grades, it is not clear whether full-day pre-K does so relative to half-day pre-K. We therefore pursue the following research questions:

1.   What is the effect of full- versus half-day pre-K offers on the likelihood of an SE designation in grades K-3, and is that effect concentrated among children who enter pre-K assessed as having potential need for SE supports?

2. Are there differential effects of full- versus half-day pre-K offers across key student subgroups, including: (a) between children whose primary home language is or is not English, and (b) between male and female students?

Background and Context

FDPK Pilot Program RCT Context

WPS has provided universal half-day pre-K since 1995—three hours per day, four days per week for a total of 12 hours weekly (see Online Appendix A for additional background information on WPS). However, in the years leading up to the study, uptake of pre-K offerings in the district had declined. Only about half of the approximately 1,100 eligible four-year-olds living in the district were enrolled in the half-day program in the 2014-15 school year (Interview With Early Childhood Department Leadership, 2016). In an effort to serve more Westminster families, WPS leveraged a PFS financing model to experiment with offering full-day programming in addition to its existing half-day offerings. WPS opened seven new full-day classrooms for four-year-olds in August 2016 and expanded to nine full-day classrooms in 2017 and 2018. Full-day classes met six hours per day, five days per week for a total of 30 hours—more than double the hours provided in half-day.[3] Because WPS anticipated that there would be more demand for full-day classes than they could initially meet, from 2016/17 to 2018/19 they used a lottery to randomly allocate limited full-day offers. Children whose families applied and met eligibility requirements (e.g., residency in Westminster, age 4 by October 1, and no preexisting SE needs that could not be met in a full-day classroom) were randomized to receive an offer of a seat in full-day (treatment, N = 399) or half-day pre-K (control, N = 396). As a result, the research team was able to leverage a longitudinal, three-cohort RCT to estimate the effects of full-day versus half-day offerings on a variety of outcomes, including subsequent SE designations.

Background on PFS Model

While the provision of SE services is an essential component of public education, mandated by federal law, it is also costly (Individuals with Disabilities Education Act [IDEA], 2004a). In fiscal year 2022, the total average per pupil cost for a student with a disability in Colorado (CO) was between $14,127 and $15,726, depending on students' specific disability classifications (U.S. Department of Commerce, U.S. Census Bureau, 2022; WestEd, 2022). Federal funding to support students with disabilities averages just 15% of the total necessary costs, leaving states and districts to find ways to cover remaining needs (Baller & Barry, 2016; Kolbe, 2019).

---

[3]While many aspects of the full- versus half-day conditions were the same (e.g., both were taught in English, same curriculum, same teacher training requirements, same student-teacher ratio), full-day classrooms had 18 more hours of class time each week. The largest differences between full- versus half-day pre-K classroom time use was that full-day classrooms provided children with a nap (whereas a nap is not part of a half-day schedule), and full-day classrooms spent almost double the amount of time in unstructured play. With regards to instruction, both classroom types allocated similar proportions of the class day to "academic" activities such as reading/literacy, math, social studies, and science. In a separate paper, the research team leverages observational data to further characterize how time was allocated differently in full- and half-day classes (Denker & Atteberry, 2024).

WPS' FDPK pilot was funded via a PFS model—a structure through which a private funder covers the front-end costs of implementing a promising but untested intervention (here, full-day pre-K). A predetermined "success" target is defined, related to potential cost savings for the district. The district then repays the funder for any "successful" outcomes—a structure which seeks to induce overall cost-savings for the district. If the intervention does not have the intended effects, the district incurs no costs. PFS models therefore seek to shift the risk of trying new interventions from public agencies to front-end payers who, in turn, are incentivized to support interventions with a strong evidentiary basis for success (Adatto & Brest, 2020).

As stated in the Introduction, WPS identified its PFS "success" target as reducing SE designations, and in doing so were thoughtful about both ways in which full-day pre-K might accomplish this in a way that was beneficial for students. With the costs of SE increasing annually (Chambers et al., 2004; Lieberman, 2023), the district hoped that full-day pre-K could beneficially reduce SE designations in later grades. It is worth noting, however, that simpler PFS determination formulas were specified before the researchers joined the project, and the estimated effects presented herein did not determine payouts.

Conceptual Framework

An underlying complexity of this analysis is that it is unclear whether full-day pre-K would cause SE designations to increase or decrease in elementary grades and, furthermore, whether those changes would be desirable. To address this complexity, we next articulate a framework for thinking about potential theories of change wherein *either* increasing or decreasing SE designations would be beneficial for students, as depicted in Figure 1.
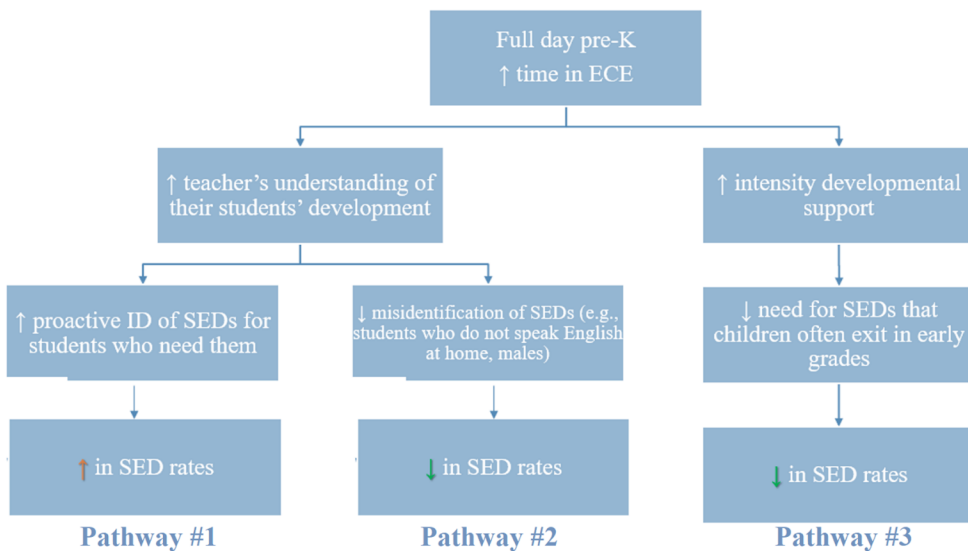
Figure 1. Theories of Change for Benefits of Full-Day Pre-K Presenting as Either Increases or Reductions in SE Designation Rates.
Note. This figure presents the theory of change for positive effects of full-day pre-kindergarten to be observed as either increases or reductions in special education designation rates. Abbreviations: Pre-K = pre-kindergarten; ECE = early childhood education; ID = identification; SEDs = special education designations.

How Could Beneficial Full-Day Pre-K Effects Present as Increased SE Designation Rates?

If SE Designations Connect Students to SE Services That Benefit Them

If SE designations trigger students to receive SE services and those services improve students' outcomes, it would be a positive result for full-day pre-K to cause *more* students who would benefit from these services to become SE designated (Pathway # 1 in Figure 1). We therefore consider the viability of this pathway by first weighing the existing evidence on whether SE designations and services are causally linked to improved outcomes for students.

This question is difficult to answer for at least two fundamental reasons: First, the "true" number of students who should be identified as having an SE designation is unknown. Second, it is empirically challenging to capture whether SE services definitively lead to better outcomes, since observational comparisons between students with and without SE designations are fraught with potential bias. Researchers have used various quasi-experimental methods to address this, with mixed results.

One approach, student fixed effects, examines changes in individual students over time. Studies using this method (Hanushek et al., 2002; Hurwitz et al., 2020; Schwartz et al., 2021) generally find SE services improve academic outcomes, especially in math and for younger students. However, this method is limited to students whose SE status changes within the study period. Another method, propensity-score matching, attempts to compare similar students with and without SE services. Findings are mixed—some studies (Morgan et al., 2010; Sullivan & Field, 2013) report lower academic scores for SE students but improvements in learning behaviors. However, this approach relies on the assumption that all relevant factors are accounted for, which is difficult to ensure. A third approach, an instrumental variables framework, leverages endogenous variation via policy changes to estimate SE's long-term effects. Ballis and Heath (2021) found that reduced SE access in Texas led to lower high school completion and college enrollment rates, suggesting SE services have positive long-term benefits. Overall, while findings are mixed across methodological approaches, evidence largely suggests SE services are effective in improving academic outcomes for students.

How Could Beneficial Full-Day Pre-K Effects Present as Reduced SE Designation Rates?

If a Greater Dosage of Pre-K Better Prepares Students and Reduces Need for SE Services

There are two paths through which a beneficial effect of full-day pre-K would manifest as an overall *reduction* in SE designation rates (Pathways # 2 and # 3 in Figure 1). First, Pathway # 3 posits that access to a more intense "dosage" of pre-K may lead students to enter elementary grades with a lower need for SE services, if the additional time in class is used to address gaps in early learning. We next consider existing evidence on whether this seems to be the case.

While we are unaware of any existing research linking pre-K intensity (full/half-day) and SE designations, there is research documenting the effects of pre-K broadly on SE placement. A meta-analysis by McCoy et al. (2017) estimates that some

(i.e., either full-time or part-time) pre-K participation, relative to no pre-K participation, results in an average decrease in SE placements of 8.1 percentage points (p<.01). Another meta-analysis focused on SE referral differences across genders further finds that boys experience lower rates of SE referrals following participation in pre-K, up to 9.5 years after students' pre-K participation (Magnuson et al., 2016).

Individual studies have directly assessed the effects of pre-K participation on the likelihood of SE placement across several contexts. Lipsey et al. (2018) study the longer-term effect of Tennessee's Voluntary Pre-K Program (VPK), relative to no pre-K, on participants' SE placement. They find that, among the sample of students randomly assigned to VPK and who ultimately attended, a higher proportion of VPK children have an Individualized Education Program (IEP) in later grades than control group children (those who were not randomly assigned to VPK and who did not attend). This finding conflicts with prior research on the long-term impacts of North Carolina's Smart Start and More at Four early childhood education (ECE) and pre-K programs and New Jersey's Abbott Preschool program, which conclude that participation in these programs results in an average reduction in SE placement (Bai et al., 2020; Barnett & Jung, 2021). If the provision of a higher dosage of pre-K operates similarly to the effect of attending any pre-K, then we might expect full-, relative to half-day, pre-K to better address early developmental needs of children and allow some to avoid the need for SE designations as they enter K-12 settings.

If a Greater Dosage of Pre-K Reduces SE Overidentification for English Learners and Boys

The other path through which we hypothesize that positive full-day pre-K effects could manifest as a *reduction* in SE designations is through avoiding the misidentification of students for SE among populations overrepresented therein (Pathway # 2 in Figure 1). Disproportionate representation of certain groups of children in SE has been an area of concern since Dunn's (1968) seminal article over 50 years ago, and a growing body of research shows both students who are learning English as a non-primary language[4] and male students continue to be overrepresented in SE today when differences in language skills and differences in developmental timing by gender are conflated with developmental deficits.

We first turn to research on the potential for young language learners to be overrepresented in SE often due to their developmentally appropriate language behaviors being misidentified with a developmental delay. Samson and Lesaux (2009) find that while language-minority learners are nationally *underrepresented* in SE in grades K and 1, they are overrepresented in grade 3 across all disability categories. They also find that SE designations are significantly predicted by both a child's language status, as well as teachers' ratings of students' English language and literacy skills.

---

[4]Varied terminology is used to describe students' language acquisition status (e.g., emergent bilinguals, English language learners, language minority learners, non-English proficient learners.) In the current study, children's pre-K language status comes from a question on the family pre-K application, phrased as: "What is the primary language spoken at home?". In the current study, we therefore refer to students' primary/non-primary home language, though when referring to extant literature, we use the terminology used by the authors.

Research over the past two decades examining students in California, Texas, and another southwestern state has highlighted consistent patterns of overrepresentation in SE among students whose home language is not English (Artiles et al., 2005; Linn & Hemmer, 2011; Sullivan, 2011)

Research on children's developmental trajectories by gender consistently shows earlier learning development among girls at young ages. For example, girls often develop language and fine motor skills earlier than boys (Huttenlocher et al., 1991; Piek & Edwards, 1997). Since these are considered typical variations in developmental timing rather than indicators of true developmental delays, this can lead to the overidentification of boys for SE in early years. Using national data, Coutinho and Oswald (2005) find that boys are 1.33 times more likely than girls to be identified with an intellectual disability, 2.04 times more likely to be identified with a learning disability, and 3.43 times more likely to be identified with a serious emotional disturbance. Hibel et al. (2010) similarly leverage nationally representative early childhood longitudinal study data and find that boys are significantly more likely than girls to be placed in SE by grade 5. And, in one district-level study, Sullivan and Bal (2013) find the highest level of risk for disability identification across subgroups to be among males and students from low-income backgrounds.

We have thus outlined competing, evidence-based theories of change about whether desirable full-day pre-K effects would manifest as SE designation rate increases or decreases. We cannot entirely disentangle these pathways in the current study, since doing so would require knowledge of an unobservable—whether each student would have indeed benefited from SE services had they been designated to receive them. However, we can establish that full-day pre-K reduces SE designations compared to half-day—a finding inconsistent with Pathway # 1 in Figure 1. We also examine whether reductions are concentrated among subgroups historically overrepresented in SE, which would support Pathway # 2 as the likely mechanism. We revisit our conceptual framework in the discussion, in light of our findings.

Data, Analytic Sample, and Analytic Approach

Data Sources

This article utilizes data from three sources: (1) administrative data on students from WPS, (2) primary baseline assessment data collected by the research team during the initial pre-K year, and (3) administrative data from the Colorado Department of Education (CDE) tracking each study child in grades K-3 across CO. For (1), in addition to typical baseline demographic student-level data, WPS required all eligible and interested families to complete a WPS General Preschool Application when they applied for pre-K, providing a rich set of pretreatment descriptive data on all children in the sample (see Table 1). They also provided SE designation data in grades K through 3 for study students who remained in the district. For (2), the research team administered baseline assessments to all study children on the PPVT (to measure receptive vocabulary skills in English) and the ESI-R, an in-depth screener for developmental delays in the domains of language and cognition, visual motor/adaptive skills, and gross motor skills, administered in

the child's primary language. These baseline assessments were conducted by an external team of trained assessors and were not shared with WPS. WPS uses its own, separate processes to ultimately determine SE designations (see Online Appendix B for additional background on both assessments). For (3), the CDE provided binary SE designation status indicators for study children who had moved to other districts in CO.

Analytic Sample

A total of 795 children were part of the FDPK study (399 randomly offered full day, 396 randomly offered half day). Table 1 shows that this RCT sample largely consisted of low socioeconomic status (SES) families—86% qualified for the free or reduced-price lunch program, and 66% of parents reported their highest level of education as a high school degree or less. Roughly 70% of the overall sample identified as Hispanic, and 45% of families reported that English was not the primary language spoken at home. At the time they applied to WPS for pre-K, 20% of families indicated their child had low language development and 46% indicated their child had low social development. About 14% indicated that an immediate family member had a history of special needs. On baseline assessments, the average (fall of pre-K) PPVT score was 91.5, with a standard deviation (SD) of 18.7, with 80% of the RCT sample scoring between 67

Table 1. Full study sample descriptive statistics.

	Full Study Sample			
	Mean	SD	N	% Missing
Baseline Demographics				
Race/ethnicity Hispanic	0.70	–	795	0.0%
Race/ethnicity White	0.19	–	795	0.0%
Race/ethnicity other/unknown[a]	0.12	–	795	0.0%
Home language not English	0.45	–	774	2.6%
Parent education ≤ high school	0.66	–	793	0.3%
Eligible for free lunch program	0.72	–	795	0.0%
Eligible for reduced-price lunch program	0.14	–	795	0.0%
Female	0.49	–	795	0.0%
With family history of special needs	0.14	–	793	0.3%
With low language development	0.20	–	793	0.3%
With low social development	0.46	–	793	0.3%
Child's age (in years)	4.51	0.29	795	0.0%
Assessment Variables				
PPVT age-adjusted baseline (fall of PK) score	91.49	18.74	711	10.6%
ESI-R baseline (fall of PK) total score	19.53	5.87	729	8.3%
SE Designation Outcome Variables				
Has a SE designation at the end of Kindergarten	0.11	–	757	4.8%
Has a SE designation at the end of 1st grade	0.11	–	742	6.7%
Has a SE designation at the end of 2nd grade	0.13	–	736	7.4%
Has a SE designation at the end of 3rd grade	0.14	–	721	9.3%

Note. This table presents descriptive statistics for the full study sample on variables collected by Westminster Public Schools (WPS) or the study team. Special education (SE) designation outcome variables for children who did not attend WPS in elementary school are from the Colorado Department of Education. 795 study children were randomized. Demographic and family history questions come from the general application for WPS preschool. To compare the study sample to WPS overall, we examine demographics from the Common Core of Data from NCES for 2016-17. The full district is 49% female, 70% free lunch eligible, 11% reduced lunch eligible, 77% Hispanic, and 1% Black (U.S. Department of Education and National Center for Education Statistics, 2021). PK = Preschool; PPVT = Peabody Picture Vocabulary Test; ESI-R = Early Screening Inventory-Revised. [a] Other = Asian, Black, Pacific Islander, American Indian/Alaskan Native.

and 115.[5] The average baseline ESI-R score was about 19.5 (median of 20, SD of 5.9), which is close to the cutoff between "rescreen" and "OK".[6] In primary analyses, we compare findings among students who scored in the bottom 50% of the baseline ESI-R distribution (includes all children that would be categorized as "Refer" and 80% of those categorized as "Rescreen") to those in the top half of that distribution (95% of whom would be categorized as "Ok").

Outcome Variable

The primary outcome variable in this study is a binary indicator of whether or not a student had an SE designation in each elementary grade K, 1, 2, or 3.[7] The process by which WPS students receive an SE designation follows federal "child find" requirements and common practices for student identification nationwide (IDEA, 2004b). After evaluation exams are administered, the SE team determines if a student will be designated as having SE needs or not, and if so, under which disability classification. For all SE-designated students, school personnel work with parents and teachers to develop an IEP that identifies the supports that will be provided to the student and the student's academic goals for the upcoming year. IEPs are reevaluated annually for adjustments or discontinuations.

Table 1 shows that between 11% and 14% of the RCT sample had an SE designation in each early elementary grade. Although a small number of students attrit from the analytic sample by leaving the state by grade 3, Table 1 underscores that the rates of missing outcome data remain low across the study period, starting from 4.8% in K and increasing only to 9.3% by grade 3. For WPS-enrolled students, we also observe the primary disability *type* for each student's SE designation (Table 2; see Online Appendix Table D1 for primary disability type by study cohort). Consistent with national trends, the most common disability types across all grades are speech/language impairment, developmental delay, and specific learning disability. These three disability types comprise nearly two-thirds of students with SE designations in WPS. Ideally, our analysis would concentrate exclusively on types of SE designations that might plausibly be influenced by attending full- versus half-day pre-K (for instance, ECE interventions are unlikely to result in an exit from SE designations for visual impairment, deaf/blindness, or hearing impairment classifications, but classifications of specific learning disability and developmental delay are theoretically responsive to changes in educational experiences). However, data limitations for disability type constrain our

---

[5]The nationally-normed mean standard score on the PPVT is 100 and the standard deviation is 15. Student scores between 85 and 115 are considered within the average range, on a total scale of 20 to 160 (Dunn & Dunn, 2013).

[6]On the ESI-R, each child receives a summary score between 0 and about 36. Scores are then translated into one of three categories, normed to a child's age: 'Refer,' indicating that a child may need SE services to support a learning condition and should be referred for an in-depth assessment; 'Rescreen,' indicating that a repeat ESI-R should be administered within the next 1-2 months; or 'OK,' indicating that the child appears to be developing without any delays (Meisels et al., 1993). For children aged 4.0 to 4.5 years old, the cutoff between rescreen is 21 (English version) and 19 (Spanish version).

[7]The research team received annual SE designation indicators from both WPS and CDE, and for students with data in both sources, the agreement rate was 97%. In those rare cases where SE designation status conflicted, we privileged the WPS-provided data.

Table 2. Primary SE disability type by grade.

	All Cohorts			
	K	1	2	3
# Missing Data on SE Designation Status	38	53	59	74
	(5%)	(7%)	(7%)	(9%)
# Does Not Have a SE Designation in Grade	673	659	642	621
	(85%)	(83%)	(81%)	(78%)
# Has a SE Designation in Grade	**84**	**83**	**94**	**100**
	(11%)	(10%)	(12%)	(13%)
Total Study Students	795	795	795	795
	(100%)	(100%)	(100%)	(100%)
Among Students with SE Designations				
Speech or Language Impairment	32%	27%	23%	15%
Developmental Delay	33%	25%	18%	9%
Specific Learning Disability	2%	6%	5%	16%
Serious Emotional Disability	4%	5%	4%	4%
Intellectual Disability	1%	1%	2%	1%
Autism Spectrum Disorders	4%	4%	2%	2%
Hearing Impairment	2%	2%	2%	2%
Other Health Impairment	0%	1%	1%	2%
Multiple Disabilities	0%	0%	0%	1%
Type Not Provided	21%	29%	41%	48%
Total Study Students with a SE Designation	**84**	**83**	**94**	**100**
	(100%)	(100%)	(100%)	(100%)

Note. The following designations exist but do not appear in our data: visual impairment, deaf/blindness, traumatic brain injury, orthopedic impairment. SE = special education. Data on SE designation status is from Westminster Public Schools (WPS) and the Colorado Department of Education (CDE). Data on SE designation type is from WPS. We do not have SE designation type information for a subset of the sample because we only receive this data from WPS and not from CDE.

analysis to assessments of full-day pre-K impacts on the simpler binary (yes/no) outcome of whether or not each child has any SE designation in each grade. For that outcome, not only is the overall missingness rate low, Table 3 shows that missing data rates across treatment and control groups are very similar—within a percentage point of one another across grades (see Online Appendix Table D2). We nonetheless explore the sensitivity of our results through assumptions about the nature of this missing data in subsequent analyses.

RCT Implementation

Baseline Covariate Balance

Due to randomization, the treatment and control groups should be similar at baseline, though it is possible for chance imbalances to occur. To confirm, we examined the set of baseline covariates for observable differences between treatment groups by fitting a series of models in which each covariate was regressed on an indicator for whether the family was randomly offered full- or half-day pre-K. These models, like those in our main analyses, include a set of lottery fixed effects for family's first-choice school site and each round of the lottery. Each row of Table 4 presents a separate regression with the same independent variable (treatment status) but a different baseline covariate outcome variable (logistic regression was used for binary outcomes). Results are shown in both standardized difference metrics (Cohen's *d*) and original metrics (e.g., percentages, mean scores) and illustrate that the full- and half-day groups were similar at baseline across all covariates. We also observe no systematic patterns of advantage

Table 3. Count and percent of study students missing SE designation outcome information, by grade and treatment status.

	Kindergarten		Grade 1		Grade 2		Grade 3	
	Full-Day (T)	Half-Day (C)	Full-Day (T)	Half-Day (C)	Full-Day (T)	Half-Day (C)	Full-Day (T)	Half-Day (C)
	All Study Children							
# Missing Data on SE Designation Status	18	20	24	29	30	29	39	35
	(5%)	(5%)	(6%)	(7%)	(8%)	(7%)	(10%)	(9%)
# Not Missing Data on SE Designation Status	381	376	375	367	369	367	360	361
	(95%)	(95%)	(94%)	(93%)	(92%)	(93%)	(90%)	(91%)
Total Number in T/C Sample	399	396	399	396	399	396	399	396
	(100%)	(100%)	(100%)	(100%)	(100%)	(100%)	(100%)	(100%)

Note. SE = special education. Data on SE designation status is from Westminster Public Schools and the Colorado Department of Education.

Table 4. Baseline covariate balance, full RCT sample.

Pretreatment covariate	Treatment M	Control M	Raw difference	Cohen's d	T/Z statistic	p		N
% Race/ethnicity Hispanic	68.2%	71.0%	−2.8%	−0.05	−0.70	0.485	−	795
% Race/ethnicity White	20.8%	16.4%	4.4%	0.09	1.26	0.208	−	795
% Race/ethnicity other/unknown	11.0%	12.6%	−1.6%	−0.04	−0.54	0.590	−	795
% Home language not English	43.2%	47.1%	−3.9%	−0.06	−0.82	0.414	−	774
% Parent education ≤ high school	65.9%	66.5%	−0.6%	−0.03	−0.40	0.686	−	793
% Eligible for free lunch program	73.9%	70.7%	3.2%	0.06	0.80	0.423	−	795
% Eligible for reduced-price lunch program	13.5%	14.6%	−1.1%	−0.04	−0.49	0.621	−	795
% Female	48.1%	49.7%	−1.6%	0.00	−0.04	0.965	−	795
% With family history of special needs	16.5%	11.7%	4.9%	0.12	1.67	0.095	−	793
% With low language development	21.8%	18.0%	3.8%	0.05	0.67	0.505	−	793
% With low social development	43.4%	48.5%	−5.1%	−0.04	−0.52	0.601	−	793
Child's age (in years)	4.50	4.52	−0.02	−0.07	−0.91	0.362	−	795
PPVT age-adjusted baseline (fall of PK) score	92.14	90.80	1.34	0.09	1.20	0.229	−	711
ESI-R baseline (fall of PK) total score	19.74	19.30	0.44	0.02	0.31	0.755	−	729

Note. We use logistic regression for binary outcomes. Cohort lottery-round first choice school fixed effects are included to appropriately account for how the randomization was conducted. For a few outcomes (here, baseline covariates) there were cohort lottery-round first choice school sites that had no variation in the covariate outcome (e.g., none of the students who applied to a given school in a given round of the lottery in a given cohort were eligible for the reduced-price lunch program). N's in the table represent the number of non-missing observations of the outcome; however, sites with no variation in the outcome do not contribute to the estimate of the mean difference in the covariate between the treatment and control groups. Demographic and family history questions come from the general application for Westminster Public Schools preschool. Parents were asked to answer yes or no to the following questions: Q1: Has an immediate family member received special education services? Q2: Is your child in need of language development including, but not limited to, the ability to speak English? Q3: Does your child have problems with social situations? PK = Preschool; PPVT = Peabody Picture Vocabulary Test; ESI-R = Early Screening Inventory-Revised (− for not sig (p > 0.10), + for p<.10, * for p<.05, ** for p<.01, and *** for p<.001).

or disadvantage between the groups—additional evidence of an overall balanced distribution. Given the focus of the current study, we might be particularly concerned if there was evidence that baseline characteristics that might connect tightly with future propensity to have an SE designation tended to be higher in the treatment group. This does not appear to be the case. If anything, a slightly higher percentage of treatment children reported a family history of special needs, and the treatment group had a very slightly higher baseline ESI-R average score.

The main threat to internal validity in a longitudinal RCT is differential attrition from the treatment and control groups in ways that introduce subsequent imbalances

across groups over time. Though missing outcome data rates are low, and treatment/control rates stay within one percentage point of one another over time, we repeat the previous baseline covariate balance exercise, comparing results in the full sample (Table 4) to the non-attrited sample (see Table 5). For parsimony, we present the results for the sample of all study participants in 3rd grade—those with the greatest opportunity for differential attrition to emerge over time—but results for the samples in grades K, 1, and 2 appear in Online Appendix Tables D3–D5. Across the 56 hypothesis tests conducted (14 baseline covariates each in the four grade samples), only one exhibited a statistically significant difference (p<.05)—fewer than we would expect by chance alone. In sum, we do not find evidence of differential attrition across grades over time.

Take-up of Pre-K Offers

While families were randomly assigned to *offers* of full- or half-day pre-K slots in WPS, ultimate decisions on whether to accept and participate were up to families. This has no bearing on the validity of the ITT effect estimates presented in our main results, however it is relevant to the interpretation of the TOT estimates presented in Online Appendix C, and is an important contextual feature of the study. We briefly discuss take-up patterns for the interested reader.

There was a strong relationship between the randomized full- vs. half-day offers and whether or not a child actually attended full-day (depicted visually in Online Appendix Figure C1). Among the 399 families randomly offered full-day spots, the vast majority (N = 357, 89.5%) then attended full-day pre-K in WPS, while only 10.5% did not. Among the 396 families randomly offered half-day spots, the vast majority (N = 358, 90.4%) indeed did not attend full-day pre-K, while only 9.6% attended full-day anyway (more on this crossover below). For more on how all RCT families covered their childcare needs throughout the workweek, regardless of whether they enrolled in WPS pre-K or not, please see Grover, Malhotra, & Atteberry, 2025 (working article). For a more detailed discussion of take-up patterns and implications for interpreting TOT estimates, see Online Appendix C.

A small portion of study participants crossed over between treatment groups: 2.5% of 399 families offered full-day switched to half-day in WPS (N = 10), and 9.6% of students who were offered half-day ultimately spent time in a full-day class (N = 38). This was made possible by district requirements to have 16 students per pre-K classroom and to fill any vacancies that arose throughout the school year. Anticipating this, the research team selected half-day RCT students to fill the seats in full-day classrooms using a randomized waitlist.[8]

---

[8]In analyses conducted elsewhere, the research team estimated baseline covariate group mean differences between those who do and do not crossover, among all crossovers, and separately among those assigned to treatment and control conditions. Of the 39 hypothesis tests of mean differences at baseline in those analyses, only one was statistically significant; in the treatment group, those who crossed from full- to half-day had slightly higher baseline PPVT scores. While some of the mean differences between those who did and did not crossover in the treatment group at first appeared large, because only 3% crossed from full- to half-day—a total of 13 students—differences were never statistically significant.

Table 5. Baseline covariate balance comparison: full sample vs. sample with non-missing SE designation outcomes.

Pretreatment covariate	Treatment Group				Control Group				Cohen's d and sig	
	Full Sample		Non-attritted		Full Sample		Non-attritted		Full Sample	Non-attritted Sample
	Mean	Count	Mean	Count	Mean	Count	Mean	Count		
% Race/ethnicity Hispanic	68.2%	399	69.4%	360	71.0%	396	72.3%	361	-0.051	-0.063 –
% Race/ethnicity White	20.8%	399	20.3%	360	16.4%	396	14.4%	361	0.092	0.153 *
% Race/ethnicity other/unknown	11.0%	399	10.3%	360	12.6%	396	13.3%	361	-0.038	-0.094 –
% Home language not English	43.2%	396	43.7%	357	47.1%	378	47.8%	343	-0.059	-0.074 –
% Parent education≤high school	65.9%	399	66.7%	360	66.5%	394	66.6%	359	-0.029	-0.015 –
% Eligible for free lunch program	73.9%	399	74.2%	360	70.7%	396	71.2%	361	0.058	0.041 –
% Eligible for reduced-price lunch program	13.5%	399	13.9%	360	14.6%	396	14.7%	361	-0.036	-0.023 –
% Female	48.1%	399	47.5%	360	49.7%	396	49.6%	361	-0.003	0.010 –
% With family history of special needs	16.5%	399	16.9%	360	11.7%	394	11.1%	359	0.120	0.147 –
% With low language development	21.8%	399	22.2%	360	18.0%	394	16.4%	359	0.047	0.094 –
% With low social development	43.4%	399	44.2%	360	48.5%	394	47.9%	359	-0.037	-0.008 –
Child's age (in years)	4.5	399	4.5	360	4.5	396	4.5	361	-0.068	-0.063 –
PPVT age-adjusted baseline (fall of PK) score	92.1	367	92.4	333	90.8	344	90.4	319	0.095	0.129 –
ESI-R baseline (fall of PK) total score	19.7	375	19.8	341	19.3	354	19.2	325	0.024	0.052 –

Note. We use logistic regression for binary outcomes. "Full Sample" refers to all study participants originally assigned to each treatment group. The "Non-attritted" sample refers to study participants with special education designation information at the end of 3rd grade. PK = Preschool; PPVT = Peabody Picture Vocabulary Test; ESI-R = Early Screening Inventory-Revised (– for not sig (p > 0.10), + for p<.10, * for p<.05, ** for p<.01, and *** for p<.001).

Analytic Approach

The study employs a RCT with a blocked design, where randomization was conducted within each family's first-choice school site (i.e., within "lotteries"). This blocking approach ensured that families' site preferences were accommodated while also reducing the likelihood of chance imbalances between treatment and control groups. We estimate the ITT effect of being offered a full-day pre-K seat (compared to a half-day seat) on SE designations in grades K through 3. Because offers were randomly assigned, variation in treatment status is exogenous, allowing for straightforward causal inference. (Again, for additional discussion and presentation of TOT estimates, refer to Online Appendix C). The statistical model used to estimate the causal effect of an offer of full-day pre-K on early SE designation is shown in Eq. 1 (illustrated for an SE designation in K, though the same model applies to each grade level):

$$
SEDesigK_{ij} = \beta_0 + \beta_1 \left(OfferFullPK_{ij} \right) + \beta_2 \left(LowESIR_{ij} \right) + \beta_3 \left(OfferFullPK_{ij} \times LowESIR_{ij} \right) \\
+ X_{ij}\beta + \alpha_j + \epsilon_{ij}
\tag{1}
$$

The outcome, $SEDesigK_{ij}$, indicates whether or not child i in lottery j had an SE designation in K (or 1st, 2nd, or 3rd grade). Since the outcome is binary, we implement Eq. (1) using a probit specification within the generalized linear model framework. The treatment variable, $OfferFullPK_{ij}$, is a binary indicator for whether a child was randomly offered a full-day seat (1 = treatment) or a half-day seat (0 = control). We interact the treatment variable with a binary indicator, $LowESIR_{ij}$, set to 1 if the child scored in the lower 50% of the baseline ESI-R score distribution upon study entry, and 0 if in the upper 50%.[9] We also illustrate results using a continuous version of baseline ESI-R scores. Models include lottery fixed effects (α_j), as well as a vector of time-invariant and time-varying controls measured at baseline, X_{ij} (the variables in Table 4), to improve precision. The coefficient β_1 captures the ITT effect among children in the top half of the baseline ESI-R distribution, for whom we do not expect any impact on SE designations. The primary effect of interest is captured by $\beta_1 + \beta_3$, which together capture the impact of full-day pre-K offers on SE designations for children in the lower half of the ESI-R distribution—those who entered pre-K with a greater likelihood of benefiting from additional developmental support. To aid interpretation, we report both probit coefficients and marginal probability estimates for both the low and high ESI-R groups. Additionally, we assess the robustness of our findings by examining potential implications of 5% to 9% missing SE designation data. Finally, we explore potential effect heterogeneity based on populations likely to be misidentified for SE services; we do so by replacing the baseline ESI-R group interaction variable in Eq. (1) either with an indicator of home language (1 = not English) or gender (1 = girls, versus boys).

---

[9]In Online Appendix C, as an alternative to an interaction framework, we estimate ITT effects for the low versus high baseline ESI-R groups by simply running a treatment effect model separately in the two subsamples and comparing coefficients on treatment variable, $OfferFullPK_{ij}$. Results are quite similar.

Results

Comparing Overall Effects in the Bottom and Top Half of Baseline ESI-R Scores

In Table 6, we present our main results—the impact of full- versus half-day pre-K offers on the likelihood of SE designations in grades K through 3, and whether that impact depends on (interacts with) ESI-R baseline score groups.[10] Estimated coefficients for the key predictors are presented in the upper panels in original probit units, estimated marginal probabilities are presented in the lower two panels, and grade-specific outcomes are shown from left to right.

　　Among children who scored in the lower half of the ESI-R at baseline, full-day pre-K offers reduced the probability of receiving an SE designation by 5.1 to 11.2 percentage-points, depending on the grade. These represent substantial effects relative to base rates. For instance, an 11 percentage-point decrease—from 23.5% among half-day students to 12.4% among full-day students—effectively halved the probability of an SE designation in first grade for students most likely to need SE services. Estimated reductions are consistently negative and sizable across all four grades, with

Table 6. Effects of full- vs. half-day pre-K offers on SE designations, by top/bottom half of baseline ESI-R distribution.

	Grade K		Grade 1		Grade 2		Grade 3	
ITT Analysis - GLM (Coefficients in Probits)								
Offered Full-Day (Effect Among Upper 50% of ESI-R)	0.13		0.01		0.18		0.09	
(p-value)	(0.612)		(0.971)		(0.452)		(0.714)	
Indicator: Lower 50% of ESI-R	0.94	***	0.90	***	0.98	***	0.79	***
(p-value)	(0.000)		(0.000)		(0.000)		(0.001)	
Interaction: Full-Day X Lower 50% of ESI-R	−0.47		−0.55	+	−0.65	*	−0.31	
(p-value)	(0.136)		(0.077)		(0.031)		(0.291)	
Constant	−1.90	***	−1.82	***	−1.78	***	−1.70	***
(p-value)	(0.000)		(0.000)		(0.000)		(0.000)	
N	677		669		646		660	
ITT Analysis - Marginal Probabilities								
For Bottom 50% of Baseline ESI-R								
Average Probability for Half-Day Students	22.6%	***	23.5%	***	26.1%	***	23.6%	***
Average Probability for Full-Day Students	15.2%	***	12.4%	***	15.4%	***	18.5%	***
Average Difference in Pr(SE Designation)	−7.5%	+	−11.2%	**	−10.7%	**	−5.1%	
For Top 50% of Baseline ESI-R								
Average Probability for Half-Day Students	5.0%	**	6.8%	***	6.7%	***	7.9%	***
Average Probability for Full-Day Students	6.3%	***	6.9%	***	9.0%	***	9.1%	***
Average Difference in Pr(SE Designation)	1.3%		0.1%		2.2%		1.1%	

Note. In upper panel, estimated coefficients are shown in probits (p-values in parentheses). In lower panel, we present estimated marginal probabilities and differences in marginal probabilities. Results are shown for a preferred model, which includes lottery fixed effects (cohort x lottery-round x first choice school), baseline child-level demographic control variables, and baseline PPVT and ESI-R scores. Models without control variables are presented in Appendix C. We include missingness dummies in cases where respondents have missing pretreatment covariates. When our model includes the main effect and interaction, LowESIRij, we do not also include baseline ESI-R scores. The omitted/reference category is a White male student not eligible for free or reduced-price lunch of average age, with average baseline PPVT scores, with parents who have an education beyond high school, whose primary home language is English, and whose family did not report low language development, low social development, or a family history of special needs. We compute the marginal effect for each observation at its observed value and then compute the average of these effects. SE = special education; ITT = intent-to-treat; GLM = generalized linear model; PPVT = Peabody Picture Vocabulary Test; ESI-R = Early Screening Inventory-Revised, administered at study baseline (entry to pre-K). + for p < .10, * for p < .05, ** for p < .01, and *** for p < .001.

---

[10]Overall effects of both the offer and participation in full- versus half-day pre-K on the likelihood of SE designations among all study students are presented in Online Appendix Table D6. However, these findings report averages across the full study sample, and do not as usefully distinguish policy impacts between those students likely to be affected from those unlikely to be affected by this change in program intensity.

statistically significant effects in all but third grade. Meanwhile, no significant SE designation effects were observed for students in the top half of the ESI-R baseline distribution. This aligns with expectations, as these students were less likely to ever be candidates for needing SE services, making treatment effects unlikely in this group.

Figure 2 conveys these findings in a slightly different way, capturing how the predicted probability of having an SE designation differs between full- and half-day pre-K groups as a function of *continuous* baseline ESI-R scores. As expected, across all grades, as baseline scores increase, the probability of having an SE designation converges to zero for both full- and half-day groups. The greatest differences in SE designation likelihood between the two groups are observed among the students with lowest baseline ESI-R scores in this sample.

Given the modest level of attrition in outcome data due to children in the RCT sample moving out of CO throughout grades K through 3, we would not anticipate a significant impact of this attrition on our results. To confirm, we conduct a sensitivity analysis using imputed values for the missing SE designation statuses in each
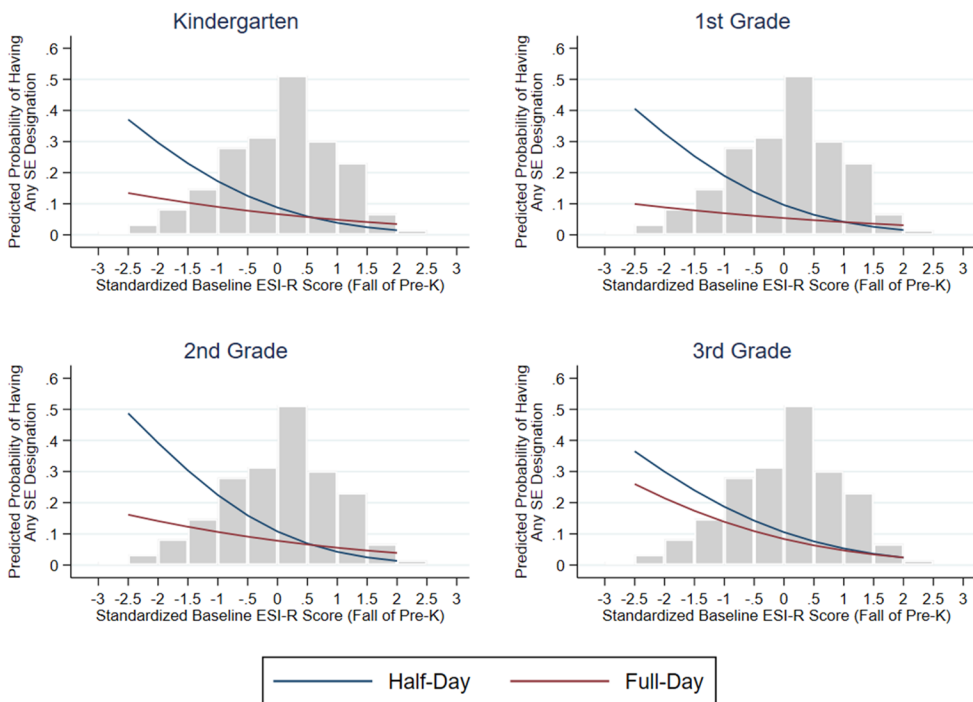

Figure 2. Relationship between Baseline ESI-R Score and Likelihood of Having an SE Designation at End of an Early Elementary Grade.

Note. This figure illustrates the relationship between students' baseline Early Screening Inventory-Revised (ESI-R) scores (standardized) and the predicted probability of having any SE designation at the end of K, 1st, 2nd, or 3rd grade, comparing relationship changes between treated students (in full-day pre-K) to control students (in half-day pre-K). Predicted probabilities for each grade are drawn from probit models (Eq. 2b in main text), which include student demographic controls, baseline Peabody Picture Vocabulary Test (PPVT) assessment scores, and fixed effects controlling for students' cohort, lottery round, and first choice school site. Assessment data come from research-team administered baseline exams while individual student SE designation data are from WPS and the CO Department of Education.

grade for the children who left the state and re-run the preceding analyses.[11] Results based on imputed outcome data are presented in Online Appendix Table D7 and show that the pattern of the results generally persists. We therefore are less concerned that estimated effects are biased due to attrition.

Overall, these results support the hypothesis that full-day pre-K reduces SE designations primarily among students who entered pre-K with some potential developmental needs, while having no impact on those less likely to require services. We next examine whether the benefits of full- vs. half-day pre-K depend on the child's home language.

Effects among Children Whose Primary Home Language is Not English

In Table 7, we find that estimated full-day treatment effects appear to be concentrated among children whose primary home language is not English. In this group, the random offer of full-day pre-K caused between 7.5 to 12.5 percentage-point reductions in SE designations, depending on the grade. For example, in grade 2: Among children whose primary home language is not English, 17.8% of those randomly offered a half-day seat had an SE designation, compared to only 5.4% of those offered full-day

Table 7. Effects of full- vs. half-day pre-K offers on SE designations, by primary home language (English or not).

	Grade K		Grade 1		Grade 2		Grade 3	
ITT Analysis - GLM (Coefficients in Probits)								
Offered Full-Day (Effect Among Home Lang English)	0.00		−0.12		0.09		0.18	
(p-value)	(0.993)		(0.509)		(0.635)		(0.302)	
Indicator: Home Lang Not English	−0.59	*	−0.61	**	−0.49	*	−0.48	*
(p-value)	(0.010)		(0.008)		(0.025)		(0.027)	
Interaction: Full-Day X Home Lang Not English	−0.57	+	−0.68	*	−0.92	**	−0.80	**
(p-value)	(0.061)		(0.030)		(0.002)		(0.005)	
Constant	−1.01	***	−0.96	***	−0.94	***	−0.98	***
(p-value)	(0.000)		(0.000)		(0.000)		(0.000)	
N	677		669		646		660	
ITT Analysis - Marginal Probabilities								
For Home Language Not English								
Average Probability for Half-Day Students	13.1%	***	14.5%	***	17.8%	***	16.8%	***
Average Probability for Full-Day Students	5.6%	**	4.2%	**	5.4%	**	7.0%	***
Average Difference in Pr(SE Designation)	−7.5%	*	−10.2%	**	−12.5%	***	−9.8%	**
For Home Language English								
Average Probability for Half-Day Students	15.1%	***	16.4%	***	16.2%	***	15.8%	***
Average Probability for Full-Day Students	15.2%	***	14.1%	***	18.0%	***	19.7%	***
Average Difference in Pr(SE Designation)	0.0%		−2.3%		1.8%		3.9%	

Note. In upper panel, estimated coefficients are shown in probits (p-values in parentheses). In lower panel, estimates are converted to marginal probabilities. Results are shown for a preferred model, which includes lottery fixed effects (cohort x lottery-round x first choice school), baseline child-level demographic control variables, and baseline PPVT and ESI-R scores. Models without control variables are presented in Appendix C. We include missingness dummies in cases where respondents have missing pretreatment covariates. We compute the marginal effect for each observation at its observed value and then compute the average of these effects. SE = special education; ITT = intent-to-treat; GLM = generalized linear model; PPVT = Peabody Picture Vocabulary Test; ESI-R = Early Screening Inventory-Revised (+ for p < .10, * for p < .05, ** for p < .01, and *** for p < .001).

---

[11]To impute missing values of the outcome (having an SE designation in K, 1st, 2nd or 3rd grade), we first specify a logistic regression model to predict designation status in each grade as a function of only baseline covariates (not including treatment status or lottery fixed effects). We use the estimated coefficients from that model to predict a probability of an SE designation for each student. For students who are missing SE designation status, we impute a value of 1 if the predicted probability is greater than 50% and a value of 0 if the predicted probability is less than 50%.

pre-K—a reduction of more than 50%. Yet we observe no apparent SE designation rate effects among children whose primary home language was English—a pattern that is consistent across all grades. The interaction coefficients are all significant, suggesting that the effect estimates for these two groups are statistically distinguishable from one another. Taken together, this suggests that indeed full-day pre-K's effect on SE designations appears concentrated among students who likely entered pre-K with less developed English language skills, which literature shows can lead to overrepresentation in SE.

Effects among Male versus Female Children

Table 8 presents results examining whether the effects of full-day pre-K on SE designation rates differ between male and female students. Here, we replace the home language interaction variable with an indicator for gender. Among female students, full-day pre-K offers led to statistically significant reductions in SE designation rates, ranging from 5.5 to 6.5 percentage points across grades ($p < 0.10$ or lower). In contrast, no significant differences in SE designation rates were observed for male students based on the randomized full- vs. half-day pre-K offers.

Notably, boys had higher overall SE designation rates than girls in grades K through 3, a pattern consistent with prior research. Young boys tend to develop certain skills at a slightly slower pace than their female peers, which may contribute to higher rates of misperceived developmental delays and subsequent SE designations. If full-day pre-K were particularly effective at reducing such gender-based misperceptions, we might expect to see that full-day reduces SE designations primarily among boys. However,

Table 8. Effects of full- vs. half-day pre-K offers on SE designations, by male vs. female children.

	Grade K		Grade 1		Grade 2		Grade 3	
ITT Analysis - GLM (Coefficients in Probits)								
Offered Full-Day (Effect Among Male Children)	0.07		−0.18		−0.09		0.11	
(p-value)	(0.706)		(0.303)		(0.607)		(0.519)	
Indicator: Female	−0.17		−0.39	*	−0.31		−0.18	
(p-value)	(0.384)		(0.049)		(0.101)		(0.341)	
Interaction: Full-Day X Female	−0.58	+	−0.28		−0.30		−0.54	+
(p-value)	(0.051)		(0.346)	+	(0.290)		(0.052)	
Constant	−1.15	***	−1.00	***	−0.95	***	−1.05	***
(p-value)	(0.000)		(0.000)		(0.000)		(0.000)	
N	677		669		646		660	
ITT Analysis - Marginal Probabilities								
For Female								
Average Probability for Half-Day Students	10.9%	***	10.1%	***	11.5%	***	12.9%	***
Average Probability for Full-Day Students	4.5%	**	4.6%	**	6.0%	**	6.5%	***
Average Difference in Pr(SE Designation)	−6.4%	*	−5.6%	+	−5.5%	+	−6.5%	+
For Male								
Average Probability for Half-Day Students	16.2%	***	19.4%	***	21.0%	***	18.8%	***
Average Probability for Full-Day Students	17.7%	***	15.3%	***	18.7%	***	21.6%	***
Average Difference in Pr(SE Designation)	1.5%		−4.1%		−2.2%		2.7%	

Note. In upper panel, estimated coefficients are shown in probits (p-values in parentheses). In lower panel, estimates are converted to marginal probabilities. Results are shown for a preferred model, which includes lottery fixed effects (cohort x lottery-round x first choice school), baseline child-level demographic control variables, and baseline PPVT and ESI-R scores. Models without control variables are presented in Appendix C. We include missingness dummies in cases where respondents have missing pretreatment covariates. We compute the marginal effect for each observation at its observed value and then compute the average of these effects. SE = special education; ITT = intent-to-treat; GLM = generalized linear model; PPVT = Peabody Picture Vocabulary Test; ESI-R = Early Screening Inventory-Revised (+ for $p < .10$, * for $p < .05$, ** for $p < .01$, and *** for $p < .001$).

our findings do not support this hypothesis, as the observed full-/half-day SE designation reductions were concentrated among girls.

Discussion

A school district's decision to offer full- versus half-day pre-K is a difficult one, partly because it is a relatively expensive policy option—essentially, doubling the number of teachers and classrooms needed to serve the same number of children who could otherwise be accommodated with an AM/PM half-day setup. At the same time, full-day pre-K could meaningfully alter children's trajectories, by increasing children's exposure to foundational experiences that support their early development and increasing the amount of time children spend with their pre-K teachers. It is unclear whether this change to the intensity of the pre-K experience leads to changes in SE designations in early elementary grades—a question which itself has both important mechanistic and financial implications.

To our knowledge, the current study is the first to provide rigorous, causal evidence on whether the offer of and participation in full- versus half-day pre-K affects students' likelihood of having an SE designation in early elementary grades. Children were randomly assigned offers of half-day pre-K (business as usual) or full-day pre-K and were followed for four years through grade 3 to track whether they had an SE designation in each subsequent grade. This tracking resulted in a longitudinal dataset that made possible the exploration of these causal effects and potential differential effects for key subgroups of students.

Before reflecting on our estimated effects, it is important to first appreciate why interpreting full-day pre-K's impact on SE designations is complex. Not only are there plausible *a priori* hypotheses for observing either increases or decreases in SE designations, but it is also unclear which directional change would represent a beneficial outcome for students. Would it be desirable for SE designation rates to increase because a more intensive pre-K experience helps more students get identified for the SE services they need? Would it instead be desirable for SE designation rates to decrease, because some students are able to have their developmental needs met before K-3 SE services become necessary? Or are lower SE designation rates desirable because more time with a full-day pre-K teacher helps counteract stereotypes and misunderstandings that lead to SE overrepresentation of certain student subgroups?

We cannot entirely adjudicate among these possibilities, because we cannot know whether each student would have benefited from SE services had they been designated to receive them. However, we can return to our theories of change shown in Figure 1 and consider which pathways are most consistent with our findings. For instance, our evidence is *inconsistent* with Figure 1's Pathway # 1, wherein full-day pre-K benefits would manifest as SE designation *increases*. This did not occur in WPS. Instead, we find that full-day pre-K seems to meaningfully *reduce* the likelihood of having an SE designation in early elementary grades and, sensibly, only among students assessed at baseline as having any possible need for developmental supports. In contrast, random full- versus half-day pre-K offers had no effect on the likelihood of SE designations among students who, at baseline, were assessed as having no potential need for SE services (shown in Figure 2).

If these SE designation reductions were a positive outcome of offering full-day pre-K, then Figure 1 offers Pathways # 2 and # 3 as possible theories for that change. We find evidence that both could be at play here. With respect to Pathway # 2, we find that SE designation reductions were concentrated among children whose primary home language was not English. Prior research demonstrates that for such children, natural and developmentally-appropriate English language acquisition can be misinterpreted as a learning disability or a developmental delay, resulting in the overrepresentation of this subgroup of students in SE (Linn & Hemmer, 2011; Sullivan, 2011). This finding supports the potential for full-day pre-K to mitigate the tendency to conflate English language acquisition with developmental disabilities.

At the same time, we did not find that full-day pre-K lowered K-3 SE designations among boys, another subgroup often overrepresented in early grade SE. Instead, we find that full-day pre-K reduced SE designations only among girls. Given the hypothesis that full-day pre-K could beneficially reduce SE designations by combatting tendencies to overidentify boys, this concentrated full-day-induced SE reduction among females is an unexpected finding. This also runs somewhat counter to prior research that finds that ECE programming broadly tends to lower rates of SE referrals for boys (Magnuson et al., 2016). Without data to support analyses on the particular classifications most responsive to this change in pre-K duration, we cannot speak to mechanisms that would further elucidate this gender-related finding, though we recommend future work consider this an important area for additional exploration.

The current study does not directly explore whether SE reduction effects operate through Pathway # 3—wherein the greater intensity of pre-K acts preventatively so that students enter elementary school with a stronger skill foundation and thus less need for developmental supports they may have otherwise needed with only a half-day experience. However, prior published work suggests this might be the case. As noted in the introduction, previous findings from the same field experiment document that these random offers of full-day pre-K produced substantial, positive effects on children's receptive vocabulary skills (0.275 standard deviations) by the end of pre-K and, among children enrolled in WPS district schools, also on teacher-reported measures of cognition, literacy, math, physical, and socioemotional development. Estimated effects on end-of-pre-K ESI-R scores were also positive and substantively large (about 0.10 SDs) but not yet statistically significant (only 1 of 3 cohorts had completed pre-K at that time). This suggests that full-day pre-K may indeed prevent future need for SE services by directly improving the skillsets with which children enter K. Research on other forms of pre-K intensity—for instance, on children who received one versus two years of preschool—also find higher scores on early elementary assessments of vocabulary, math skills, reading, and working memory (Shah et al., 2017; Wen et al., 2012). Additionally, while research documenting the effects of ECE programming broadly on SE placements in elementary grades are mixed, evidence largely finds that ECE participation reduces students' SE placements in early elementary grades (Bai et al., 2020; Lipsey et al., 2018; Magnuson et al., 2016; McCoy et al., 2017). This suggests that the effect of attending pre-K could operate similarly to the provision of a higher dosage of pre-K, potentially better addressing early developmental needs of children and allow some to avoid the need for SE designations as they enter K-12 settings. Taken together, this signals support for Pathway # 3, which posits a direct relationship between a

higher dosage of developmental experiences in pre-K and reductions in the need for SE supports in subsequent grades.

It is important to acknowledge the limitations of our study's generalizability. In many respects, WPS is similar to numerous U.S. school districts—it is of typical size, suburban, and serves a predominantly low-income student population. However, WPS differs from many districts in that it primarily serves Hispanic students, and nearly half speak a language other than English at home. This linguistic and demographic composition may influence if and how full-day pre-K affects student outcomes, relative to half-day pre-K. For instance, if increased exposure to English is a key mechanism driving the benefits of extended pre-K hours, then the effects observed here may not fully generalize to districts with fewer English learners.

That said, a substantial number of U.S. public school districts serve large Hispanic populations—approximately 1 in 5 have student bodies that are at least 25% Hispanic (Reardon et al., 2019). As a result, our findings may be particularly relevant for districts seeking to improve outcomes for this subgroup. Nonetheless, further research is needed to assess whether full-day pre-K has similar effects in districts serving different racial, ethnic, and linguistic populations. Expanding studies of full- vs. half-day pre-K to a broader range of settings will be crucial for understanding the extent to which these findings apply across diverse educational contexts.

Our findings may also be relevant to the 57% of school districts offering some form of pre-K programming but who only currently offer half-day options (Civil Rights Data Collection, 2020). This group of roughly 7,500 districts has both the same "baseline" conditions as WPS in only offering half-day programming, as well as the potential to be experiencing similar mismatches between offerings and local family needs. This set of districts could therefore look to adding additional full-day services on top of existing offerings, drawing on conclusions offered by this study.

The current study also has a few important analytic limitations. First, the lack of specific disability types for about half of the sample prevents us from exploring effects among this important dimension. Yet there are undoubtedly differences in the relationship between pre-K program intensity and the likelihood of being identified as having a mild/moderate impairment relative to a more severe disability, as well as across all disability classifications. Future research that can disaggregate effects across classifications will provide valuable insight into the relationship between pre-K and special education. Second, we cannot fully disentangle whether these treatment/control group mean reductions in SE designations were ultimately helpful or harmful to children in the long term. While we have suggestive evidence that these reductions may have followed promising pathways, we cannot empirically test this without access to elementary grade assessments that are comparable across both English language status and SE needs.

While more research is needed, the current study finds a strong relationship between pre-K program intensity and SE designations during students' early elementary years. Among children who had some potential to need developmental supports as they entered their pre-K year at baseline, full-day pre-K offers reduced the likelihood of having a subsequent SE designation in grades K-2 by between 7.5 and 11.2 percentage points—a sizeable reduction, relative to a base rate of around 20 percent. In its PFS success targets, WPS estimated an annual average cost savings to the district of about $7,000 per reduction in SE-designation. These impacts would therefore represent a

meaningful cost-savings for a district, which may help to offset the costs of restructuring pre-K to serve just one roster of full-day students with each classroom/teacher, rather than serving two half-day rosters of students with the same resources. There may be other benefits to full-day pre-K, in terms of both children's academic and social skills, educational attainment, and family employment, finances, and stress—outcomes that are the focus of future research on this field experiment. Yet as costs for SE continue to rise annually, and all schools seek ways to ensure the right students are being identified for the appropriate support services, this study highlights one potential lever for policymakers and school leaders to consider when thinking about the design of future of ECE programming.

Disclosure Statement

No potential conflict of interest was reported by the author(s).

Open Research Statements

Study and Analysis Plan Registration

The analysis plan for long-term research questions addressed in this study are registered in the Registry of Efficacy and Effectiveness Studies (ID: 15380.1v1).

Data, Code, and Materials Transparency

The materials, data, and code associated with this study are not publicly available.

Design and Analysis Reporting Guidelines

This manuscript was not accompanied by a completed copy of the JREE Randomized Trial Checklist.

Transparency Declaration

The corresponding author (the manuscript's guarantor) affirms that the manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned (and, if relevant, registered) have been explained.

Replication Statement

This manuscript reports an original study.

Open scholarship

This article has earned the Center for Open Science badges for Preregistered. The materials are openly accessible at the Registry of Efficacy and Effectiveness Studies (ID: 15380.1v1).

Funding

References

Adatto, S., & Brest, P. (2020). *The Pay for Success Handbook*. Open Textbook Library.

Artiles, A. J., Rueda, R., Salazar, J. J., & Higareda, I. (2005). Within-group diversity in minority disproportionate representation: English language learners in urban school districts. *Exceptional Children*, *71*(3), 283–300. https://doi.org/10.1177/001440290507100305

Atteberry, A., Bassok, D., & Wong, V. C. (2019). The effects of full-day prekindergarten: Experimental evidence of impacts on children's school readiness. *Educational Evaluation and Policy Analysis*, *41*(4), 537–562. https://doi.org/10.3102/0162373719872197

Bai, Y., Ladd, H. F., Muschkin, C. G., & Dodge, K. A. (2020). Long-term effects of early childhood programs through eighth grade: Do the effects fade out or grow? *Children and Youth Services Review*, *112*, 104890. https://doi.org/10.1016/j.childyouth.2020.104890

Baller, J. B., & Barry, C. L. (2016). State variation in school-based disability services financed by Medicaid. *Journal of Disability Policy Studies*, *27*(3), 148–157. https://doi.org/10.1177/1044207316637545

Ballis, B., & Heath, K. (2021). The long-run impacts of special education. *American Economic Journal: Economic Policy*, *13*(4), 72–111. https://doi.org/10.1257/pol.20190603

Barnett, W. S., & Jung, K. (2021). Effects of New Jersey's Abbott preschool program on children's achievement, grade retention, and special education through tenth grade. *Early Childhood Research Quarterly*, *56*, 248–259. https://doi.org/10.1016/j.ecresq.2021.04.001

Chambers, J., Parrish, T., & Harr, J. (2004). *What are we spending on special education services in the United States, 1999-2000?* U.S. Department of Education. https://www.air.org/sites/default/files/SEEP1-What-Are-We-Spending-On.pdf

Chetty, R., Friedman, J. N., & Rockoff, J. E. (2014). Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood. *American Economic Review*, *104*(9), 2633–2679. https://doi.org/10.1257/aer.104.9.2633

Civil Rights Data Collection. (2020). *Civil rights data collection*. https://ocrdata.ed.gov/Home

Colorado Department of Education. (2020). *School view: School and district data: Data lab*. https://www.cde.state.co.us/schoolview

Coutinho, M. J., & Oswald, D. P. (2005). State variation in gender disproportionality in special education: Findings and recommendations. *Remedial and Special Education*, *26*(1), 7–15. https://doi.org/10.1177/07419325050260010201

Denker, H., & Atteberry, A. (2024). Where has all the time gone? Describing time use in full-vs. half-day pre-Kindergarten. *Early Childhood Research Quarterly*, *68*, 235–246. https://doi.org/10.1016/j.ecresq.2024.05.007

Duncan, G. J., & Magnuson, K. (2013). Investing in preschool programs. *The Journal of Economic Perspectives: A Journal of the American Economic Association*, *27*(2), 109–132. https://doi.org/10.1257/jep.27.2.109

Dunn, L. M. (1968). Special education for the mildly retarded—Is much of it justifiable? *Exceptional Children*, *35*(1), 5–22. https://doi.org/10.1177/001440296803500101

Dunn, L., & Dunn, D. (2013). *PPVT-4 technical report*. Pearson Assessments. https://www.pearsonassessments.com/content/dam/school/global/clinical/us/assets/ppvt-4/ppvt-4-technical-report.pdf

Grover, V., Malhotra, K., & Atteberry, A. (2025). Experimental Effects of Full- vs. Half-Day Pre-K on the Home Lives of Young Children and Their Families. (Working Paper).

Halfon, N., & Hochstein, M. (2002). Life course health development: An integrated framework for developing health, policy, and research. *The Milbank Quarterly*, *80*(3), 433–479, iii. https://doi.org/10.1111/1468-0009.00019

Hanushek, E. A., Kain, J. F., & Rivkin, S. G. (2002). Inferring program effects for special populations: Does special education raise achievement for students with disabilities? *Review of Economics and Statistics*, *84*(4), 584–599. https://doi.org/10.1162/003465302760556431

Hibel, J., Farkas, G., & Morgan, P. L. (2010). Who is placed into special education? *Sociology of Education*, *83*(4), 312–332. https://doi.org/10.1177/0038040710383518

Hurwitz, S., Perry, B., Cohen, E. D., & Skiba, R. (2020). Special education and individualized academic growth: A longitudinal assessment of outcomes for students with disabilities. *American Educational Research Journal*, *57*(2), 576–611. https://doi.org/10.3102/0002831219857054

Huttenlocher, J., Haight, W., Bryk, A., Seltzer, M., & Lyons, T. (1991). Early vocabulary growth: Relation to language input and gender. *Developmental Psychology*, *27*(2), 236–248. https://doi.org/10.1037/0012-1649.27.2.236

Individuals with Disabilities Education Act [IDEA]. (2004a). Individuals with Disabilities Education Act [IDEA] regulations, part B, subpart B: Section 300.111 "Child Find". National Archives Code of Federal Regulations. https://www.ecfr.gov/current/title-34/subtitle-B/chapter-III/part-300/subpart-B/subject-group-ECFRf27988d69cd5d3a/section-300.111

Individuals with Disabilities Education Act [IDEA]. (2004b). *Pub. L. 108-446, Statute No. 20 U.S.C § 1400*. U.S. Congress. https://www.congress.gov/108/plaws/publ446/PLAW-108publ446.pdf

Interview With Early Childhood Department Leadership. (2016). Prediscussion of RCT study design with WPS ECC Leadership. *Personal Communication*.

Jackson, C. K. (2018). What do test scores miss? The importance of teacher effects on non–test score outcomes. *Journal of Political Economy*, *126*(5), 2072–2107. https://doi.org/10.1086/699018

Kolbe, T. (2019). *Funding Special Education: Charting a Path that Confronts Complexity and Crafts Coherence*. National Education Policy Center.

Lieberman, M. (2023). *"Special Education Is Getting More Expensive, Forcing Schools to Make Cuts Elsewhere." Education Week*. https://www-edweek-org.colorado.idm.oclc.org/leadership/special-education-is-getting-more-expensive-forcing-schools-to-make-cuts-elsewhere/2023/04

Linn, D., & Hemmer, L. (2011). English language learner disproportionality in special education: Implications for the scholar-practitioner. *Journal of Educational Research and Practice*, *1*(1), 70–80. https://doi.org/10.5590/Jerap.2011.01.1.06

Lipsey, M. W., Farran, D. C., & Durkin, K. (2018). Effects of the Tennessee Prekindergarten Program on children's achievement and behavior through third grade. *Early Childhood Research Quarterly*, *45*, 155–176. https://doi.org/10.1016/j.ecresq.2018.03.005

Magnuson, K. A., Kelchen, R., Duncan, G. J., Schindler, H. S., Shager, H., & Yoshikawa, H. (2016). Do the effects of early childhood education programs differ by gender? A meta-analysis. *Early Childhood Research Quarterly*, *36*, 521–536. https://doi.org/10.1016/j.ecresq.2015.12.021

McCoy, D. C., Yoshikawa, H., Ziol-Guest, K. M., Duncan, G. J., Schindler, H. S., Magnuson, K., Yang, R., Koepp, A., & Shonkoff, J. P. (2017). Impacts of early childhood education on medium- and long-term educational outcomes. *Educational Researcher (Washington, D.C.: 1972)*, *46*(8), 474–487. https://doi.org/10.3102/0013189X17737739

Meisels, S. J., Henderson, L. W., Liaw, F., Browning, K., & Ten Have, T. (1993). New evidence for the effectiveness of the early screening inventory. *Early Childhood Research Quarterly*, *8*(3), 327–346. https://doi.org/10.1016/S0885-2006(05)80071-7

Morgan, H. (2020). Misunderstood and mistreated: Students of color in special education. *Voices of Reform*, *3*(2), 71–81. https://doi.org/10.32623/3.10005

Morgan, P. L., Frisco, M. L., Farkas, G., & Hibel, J. (2010). A propensity score matching analysis of the effects of special education services. *The Journal of Special Education*, *43*(4), 236–254. https://doi.org/10.1177/0022466908323007

Piek, J. P., & Edwards, K. (1997). The identification of children with developmental coordination disorder by class teachers. *British Journal of Educational Psychology*, *67*(1), 55–67. https://doi.org/10.1111/j.2044-8279.1997.tb01227.x

Reardon, S. F., Fahle, E. M., Ho, A. D., Shear, B. R., Min, J., Kalogrides, D., & Kane, T. J. (2019). Stanford Education Data Archive (Version SEDA 2019). https://purl.stanford.edu/xt779fj2637

Samson, J. F., & Lesaux, N. K. (2009). Language-minority learners in special education: Rates and predictors of identification for services. *Journal of Learning Disabilities*, *42*(2), 148–162. https://doi.org/10.1177/0022219408326221

Schifter, L. A., Grindal, T., Schwartz, G., & Hehir, T. (2019). *Students from low-income families and special education*. PK-12 Education Report. The Century Foundation.

Schwartz, A. E., Hopkins, B. G., & Stiefel, L. (2021). The effects of special education on the academic performance of students with learning disabilities. *Journal of Policy Analysis and Management*, *40*(2), 480–520. https://doi.org/10.1002/pam.22282

Shah, H. K., Domitrovich, C. E., Morgan, N. R., Moore, J. E., Cooper, B. R., Jacobson, L., & Greenberg, M. T. (2017). One or two years of participation: Is dosage of an enhanced publicly funded preschool program associated with the academic and executive function skills of low-income children in early elementary school? *Early Childhood Research Quarterly*, *40*, 123–137. https://doi.org/10.1016/j.ecresq.2017.03.004

Sullivan, A. L. (2011). Disproportionality in special education identification and placement of English language learners. *Exceptional Children*, *77*(3), 317–334. https://doi.org/10.1177/001440291107700304

Sullivan, A. L., & Bal, A. (2013). Disproportionality in special education: Effects of individual and school variables on disability risk. *Exceptional Children*, *79*(4), 475–494. https://doi.org/10.1177/001440291307900406

Sullivan, A. L., & Field, S. (2013). Do preschool special education services make a difference in kindergarten reading and mathematics skills?: A propensity score weighting analysis. *Journal of School Psychology*, *51*(2), 243–260. https://doi.org/10.1016/j.jsp.2012.12.004

U.S. Department of Commerce, U.S. Census Bureau. (2022). *2022 Annual Survey of School System Finances* [dataset]. https://www.census.gov/library/visualizations/interactive/how-did-covid-19-affect-school-finances.html

U.S. Department of Education, National Center for Education Statistics. (2021). *Local Education Agency (School District) Universe Survey* [dataset]. https://nces.ed.gov/ccd/elsi/

Wen, X., Leow, C., Hahs-Vaughn, D. L., Korfmacher, J., & Marcus, S. M. (2012). Are two years better than one year? A propensity score analysis of the impact of Head Start program duration on children's school performance in kindergarten. *Early Childhood Research Quarterly*, *27*(4), 684–694. https://doi.org/10.1016/j.ecresq.2011.07.006

WestEd. (2022). *Report on Colorado Special Education Funding Mechanisms*. https://www.cde.state.co.us/cdedepcom/sefacspecialreport