



UNIVERSITY  
of VIRGINIA

CURRY SCHOOL and  
BATTEN SCHOOL  
EdPolicyWorks

Working Paper:

# The Effects of Full-day Pre-kindergarten: Experimental Evidence of Impacts on Children's School Readiness

*Allison Atteberry<sup>2</sup>, Daphna Bassok<sup>1</sup>, & Vivian C. Wong<sup>1</sup>*

---

This study is a randomized control trial of full- versus half-day pre-kindergarten in a school district near Denver. Four-year-old children were randomly assigned an offer of half-day (four days/week) or full-day (five days/week) pre-k that increased class time by over 600 hours. The offer of full-day pre-k produced large, positive effects on children's receptive vocabulary skills (0.267 standard deviations) by the end of pre-k. Among children enrolled in district schools, full-day participants also outperformed their peers on teacher-reported measures of cognition, literacy, math, and physical development. At kindergarten entry, children offered pre-k still outperformed peers on a widely-used measure of basic literacy. The study provides the first rigorous evidence on the impact of full-day preschool on children's school readiness skills.

---

<sup>1</sup>University of Virginia

<sup>2</sup>University of Colorado Boulder

*Updated July 2018*

EdPolicyWorks  
University of Virginia  
PO Box 400879  
Charlottesville, VA 22904

EdPolicyWorks working papers are available for comment and discussion only. They have not been peer-reviewed.

Do not cite or quote without author permission. Working paper retrieved from:  
[http://curry.virginia.edu/uploads/epw/64\\_Effects\\_Full\\_Day\\_Prekindergarten.pdf](http://curry.virginia.edu/uploads/epw/64_Effects_Full_Day_Prekindergarten.pdf)

*Acknowledgements:* This research was funded by Westminster Public School District and the Smith Richardson Foundation. We thank them for their generous support. All errors are solely attributable to the authors.

EdPolicyWorks Working Paper Series No. 64. July 2018.

Available at <http://curry.virginia.edu/edpolicyworks/wp>

Curry School of Education | Frank Batten School of Leadership and Public Policy | University of Virginia

High quality early childhood education (ECE) programs can have a profound effect on children's development while simultaneously yielding substantial social returns (Blau & Currie, 2006; Heckman, 2006; Shonkoff & Phillips, 2000; Weiland & Yoshikawa, 2013; Wong, Cook, Barnett, & Jung, 2008). Further, the benefits of ECE are most pronounced for low-income children (Weiland & Yoshikawa, 2013), Hispanic children (Gormley, 2008), and Black children (Bassok, 2010), suggesting that investments in ECE may be powerful tools for tackling early childhood achievement gaps and inequality. For these reasons, public investment in ECE has grown rapidly in the United States over the past two decades (Barnett et al., 2017).

Despite the strong evidence that early investments *can* create large and lasting benefits for children, research tracking the impacts of large-scale, present-day ECE programs has yielded mixed results. Recent reviews of the literature indicate that children who attend public preschool in the year prior to kindergarten start school significantly ahead of their peers (Weiland, 2018). However, a growing body of research also suggests that the initial benefits on children's academic skills may be short-lived, dissipating quickly as children progress through school

These findings have led to heightened interest among policy-makers and researchers in identifying specific program characteristics that promote returns on early childhood investments. In a recently-released consensus statement, a group of early childhood experts stressed the need to understand how preschool can serve as “an enduring base for future learning” and emphasized a need to unpack the particular features of preschool programs that contribute to children's development (Philips et al., 2017). Traditionally, policy-makers and researchers focused on “structural” characteristics of ECE settings such as the qualifications of educators, the class size, and the staff-child ratios. More recently, there has also been substantial interest in more process-oriented features such as teacher-child interactions, curricula and access to professional

development. Despite a growing literature on the role of these quality features, our understanding of the causal relationship between specific ECE features and child outcomes remains underdeveloped with little consensus on the features—or the combination of features—that are most critical for promoting children’s development.

One salient characteristic—hours of exposure—has garnered considerable attention as a potentially important lever for supporting children’s early learning. Between 1998 and 2010, the percentage of kindergarteners in the U.S. in full-day *kindergarten* grew rapidly from 55 to 80 percent (Bassok, Gibbs, & Latham, 2018). The percentage of preschoolers in full-day preschool also increased, albeit more slowly. In 2000, 47 percent of young children attended full-day programs. By 2016 that figure rose to 54 percent (Kena et al., 2016). In addition, the Office of Head Start proposed a new performance standard in 2015 that aimed to raise Head Start’s operating hours from 448 hours a year to at least 1,020 hours per year (*Head Start Performance Standards*, 2016). Policies that increase children’s hours of exposure in ECE settings have been motivated in part by the assumption that expanding the length of the school day will provide children with more exposure to high quality learning opportunities, which, in turn, will yield greater and longer-lasting benefits. Full-day preschool programs might also attract new families who would otherwise not enroll their children in classroom-based ECE programs because their work or school schedules conflict with part-day programs.

Currently, there is little empirical evidence about the extent to which access to full-day versus half-day preschool yields larger benefits, an important gap in the ECE literature given the relative cost of expanding the length of the preschool day. Full-day preschool expansion is expensive and has the potential to divert funds away from other ECE resources that may be more impactful in promoting children’s development.

To our knowledge, this is the first rigorous randomized control trial (RCT) about the benefits of full-day preschool on children's school readiness skills. We find that relative to offers of half-day preschool, full-day preschool produces large, positive effects on children's receptive vocabulary skills (0.267 standard deviations) in the spring of their preschool year. Among children enrolled in the public preschool program, full-day preschool also yields positive effects on children's cognition, literacy, math, and physical development. Our results also suggest that positive results are still present as children start their kindergarten year. Combined, the short-term effects on children's school readiness skills suggest the *promise* for longer-term and more meaningful impacts of full-day preschool.

This study presents results from an RCT of full versus half day pre-kindergarten (pre-k) in a school system near Denver, Colorado. Section 2 provides background about the potential benefits of intensifying children's exposure to ECE, section 3 describes the context for the current study. Section 4 describes the study design, measures, and analysis models; section 5 presents the impacts of full-day pre-k on children's outcomes at the end of pre-k as well as the beginning of kindergarten. We conclude by offering recommendations for policy-makers, and areas for future work.

## **Background**

ECE programs vary substantially with respect to the structural features (e.g. teacher education levels, ratios), process features (e.g. the quality of teacher child interactions) and importantly, their contributions to children's learning (Bassok, Fitzpatrick, Greenberg, & Loeb, 2016; Morris et al., 2018; Weiland, 2018). Improving ECE at scale will require a better understanding of which particular features are most important for program effectiveness. Towards this goal, a growing body of research has examined the effect of specific program characteristics.

For instance, recent RCTs have examined the effects of professional development for ECE teachers, as well as the impacts of specific curricula and teacher-child interactions (Araujo, Carneiro, Cruz-Aguayo, & Schady, 2016; Clements & Sarama, 2008; Early, Maxwell, Ponder, & Pan, 2017; Piasta et al., 2017). However, relatively few studies in ECE have used experimental methods to examine how structural features of ECE programs, which are the primary drivers of programs' costs, impact children's development. For example, there are no experimental studies measuring the impact of teacher education levels, teacher pay, or teacher-child ratios on children's learning in ECE settings. Similarly, few studies have provided rigorous causal evidence on the link between children's learning and the intensity of an ECE program, defined broadly to encompass both the number of years children attend a program, and the number of hours they are enrolled per week.

The most effective and rigorously evaluated ECE programs provided intensive interventions for children and their families. For instance, the Carolina Abecedarian Project, one of the most-touted ECE programs for its sizable impacts into adulthood, offered full-day preschool, five days a week, from infancy to age 5 and has been linked to positive outcomes through age 30 (Campbell et al., 2012; Campbell & Ramey, 1994). However, it is not clear the extent to which these findings were *caused* by the relatively intense exposure or whether it could be explained by other features of the Abecedarian program. For example, from infancy, Abecedarian children were exposed to rich learning environments with trained child development specialists and health and medical professionals, while children in the Abecedarian control condition stayed home, without access to similar care environments. The existing literature fails to isolate the *unique contribution* of intensive exposure.

More recent studies have focused on understanding the unique role of ECE intensity but have generally done so within the context of *kindergarten* rather than preschool settings. About 80 percent of kindergarteners attend full-day programs, a striking increase from the late seventies when roughly three quarters of kindergarteners attended half-day programs (Bassok et al., 2018; Child\_Trends\_Data\_Bank, 2015). This rapid expansion of full-day kindergarten has fostered heightened interest among policymakers and researchers in understanding how children are impacted by the longer school day. We are aware of only one study that uses random assignment to identify the impact of offering full versus half day kindergarten on children's outcomes. C. Gibbs (2014) studied over-subscribed full-day kindergarten programs in Indiana, where lotteries were used to allocate full-day slots. Comparing children within the same school, she found that children randomly-assigned to full rather than half-day kindergarten scored 0.31 standard deviations higher on a literacy assessment by the end of the kindergarten year. Benefits were even larger for Hispanic children and children who began kindergarten with relatively low literacy skills.

To date, nearly all other studies tackling this question have relied on observational data, comparing children who attended full-day programs to those who attended half-day programs after accounting, to the extent possible, for selection factors at the child, family, school, or community level (Brownell et al., 2015; Gullo, 2000; Zvoch, Reynolds, & Parker, 2008). Several of the most influential studies used the Early Childhood Longitudinal Study-Kindergarten Cohort (ECLS-K), which tracks a nationally-representative sample of children who entered kindergarten in 1998 as they progress through elementary and middle school (Cannon, Jacknowitz, & Painter, 2006; DeCicca, 2007; Lee, Burkam, Ready, Honigman, & Meisels, 2006; Votruba-Drzal, Li-Grining, & Maldonado-Carreño, 2008). These studies show that at the end of the kindergarten year, children

who attended full-day programs outperform similar children in half-day programs on a variety of outcomes. However, these differences dissipate quickly, with no significant differences by the time children are in third or fifth grade.

A recent meta-analysis of 40 studies of full-day kindergarten released between 1979 and 2009 came to similar conclusions (Cooper, Allen, Patall, & Dent, 2010). It showed that at the end of kindergarten, children who attended full-day kindergarten scored about a quarter of standard deviation higher than similar children in half-day programs, but that as children progressed through the elementary school years, these differences between groups disappeared. However, the authors of this review cautioned that the evidence-base on this question “leaves much to be desired” noting that the correlational nature of the studies does not sufficiently address issues of selection and cannot be used to draw causal conclusions.

Nevertheless, the limited experimental evidence available to date (C. Gibbs, 2014), as well as the broader correlational evidence on full-day kindergarten, do support the notion that more hours of exposure to kindergarten benefits young children’s academic skill-building, at least in the short-term. It is not clear, however, whether results from the kindergarten context generalize to preschool settings. ECE programs serve younger children with unique developmental needs. Classroom practices, routines and curricula differ between ECE and kindergarten classrooms, and the teachers guiding children’s learning oftentimes differ substantially across these contexts with respect to their education, training, and compensation (Abry, Latham, Bassok, & LoCasale-Crouch, 2015; Whitebook, Phillips, & Howes, 2014).

Unfortunately, there is only a small body of literature examining the impacts of ECE intensity on children’s development *prior to the kindergarten year*. Only one unpublished study is experimental (Robin, Frede, & Barnett, 2006). That study included 294 four-year-old children

drawn from an urban school district serving mostly low-income families who were randomly assigned to full (N = 77) or half day (N = 217) classes. The half-day program consisted of 2.5- to 3-hour classes for 41 weeks; the full-day program consisted of 8-hour classes for 45 weeks. At kindergarten, full-day program children scored significantly higher on cognitive assessments compared to those in the half-day program and continued to outperform the comparison group at first-grade. However, the full-day preschool group was more advantaged at baseline compared to their half-day counterparts, limiting the interpretability of the RCT results. For example, full-day children scored significantly higher on multiple pre-intervention assessments and their mothers worked more hours per week. Given the lack of baseline equivalence, the results from this study should be interpreted with caution.

All other studies about the link between preschool intensity and child outcomes rely on non-experimental or quasi-experimental methods and yield mixed results (Herry, Maltais, & Thompson, 2007). For instance, Reynolds et al. (2014) conducted within-school comparisons of children in 11 Child-Parent Center (CPC) education programs. The CPC is a long-standing early intervention program that provides comprehensive education and family services to families in low-income neighborhoods in Chicago. Reynolds et al. compared outcomes of children who attended full- and half-day programs within the same school and found that full-day children have better attendance and scored higher on four of the six school readiness indicators, including language, math, social-emotional development and physical health. However, the authors caution that their results may be biased because the full-day program prioritized enrollment for four-year-olds, so children in the half- and full-day programs were not equivalent with respect to age at baseline.



Several studies have leveraged large datasets and regression-adjustment approaches, to examine the association between the intensity of ECE settings and child outcomes. For instance, data from the nationally representative ECLS-K suggest that the benefits of ECE programs are more pronounced for children who start attending as toddlers or spend more hours in preschool per week (Loeb, Bridges, Bassok, Fuller, & Rumberger, 2007). On the other hand, a number of large-scale studies also indicate that children who spend more hours in center-based child care exhibited somewhat higher incidences of behavioral problems (Belsky, 2002; Vandell, Belsky, Burchinal, Steinberg, & Vandergrift, 2010).

Studies that have explored the benefits of full- and half-day programs specifically within the context of the Federal Head Start program also yield mixed results. Leow and Wen (2017) used propensity score matching approaches and 2016 Family and Child Experiences Survey (FACES) data to compare outcomes for three- and four-year-old children who attend full- and half-day Head Start. Their results did not indicate any benefits of full-day classes on five academic and social outcomes in kindergarten. In contrast, in his reanalysis of data from the Head Start Impact Study Walters (2015) examined whether heterogeneity in impacts across settings was explained by observable program characteristics. He found that Head Start centers offering full-day services produced larger impacts on children's cognitive outcomes compared to centers providing only part-day programming. However, this result may have been due to the offer of full-day services, or to other unobserved program feature related to full-day programming.

Finally, in a related set of studies, researchers have examined ECE intensity by comparing the benefits of participating in one versus two years of preschool. Using observational approaches, such as propensity score matching, researchers found that relative to children with one year of preschool, those with two years showed improved performance both at school entry, and as they

progressed through the early elementary grades (Leow & Wen, 2017; Shah et al., 2017; Wen, Leow, Hahs-Vaughn, Korfmacher, & Marcus, 2012).

Taken together, these studies suggest that more intensive dosages of exposure to ECE classrooms in the preschool setting may have meaningful and at times sustained impacts on children's development. However, the empirical evidence supporting these findings is limited in terms of methodological rigor for supporting causal inferences. Our study adds to existing literature by providing new experimental evidence about the impacts of full-day preschool on a host of short-term outcomes in a low-income, largely Latino population.

### **Study Context**

Westminster Public Schools (WPS) is a public-school district located northwest of Denver that serves approximately ten thousand students annually. While about 50 percent of WPS students perform at or above proficiency on statewide exams, there are large disparities in academic achievement between groups. For instance, while virtually 100 percent of WPS's fully English-fluent students are "proficient" in grade 3 math scores, only about 15 percent of its Not-English Proficient (NEP) students achieve proficiency status ([Westminster Public Schools TCAP Results, 2014](#)). The district has been tasked with providing educational programming to a population of students that is largely non-White (83 percent), low-income (76 percent), and non-native English speaking (34 percent). In recent years, it has struggled to overcome the systemic socio-economic barriers that inhibit the performance of these students.

In order to intensify its ECE offerings, WPS used a pay-for-success funding model and secured funding to expand its pre-k program from half-day only to also include full-day classes among four-year-old children. Prior to the 2016-2017 school year, WPS provided half-day preschool for three hours per day, four days per week. However, only about half of the district's

eligible 1,100 four-year old children actually enrolled in the district pre-k program, leading district leaders to consider how to serve more Westminster families (Interview with Early Childhood Department Leadership, 2016). WPS hypothesized that many district families did not take advantage of WPS pre-k services due to the half-day and partial week program availability, which may have conflicted with family's child care needs. In the summer of 2016, WPS launched the Full-Day Pre-K Program ("FDPK") for the 2016-2017 school year. Because the district anticipated oversubscription in its full-day classes, it held lotteries to award families with FDPK slots. Families who did not receive slots in the full-day program were offered enrollment in the business-as-usual half-day program.<sup>1</sup>

To assess the efficacy of FDPK on student and family outcomes, WPS committed to a rigorous evaluation of its initiative. WPS worked with the research team to randomly assign offers of full- and half-day pre-k to eligible families. In 2016-17, the district opened seven new full-day pre-k classrooms that were available for six hours per day, five days a week; the half-day program ran as it has in past, with classes available for three hours per day, four days a week. Compared to the half-day program, FDPK more than doubled the number of hours per week for children in ECE settings and added more than 600 classroom-hours over the school year. In part, these additional hours were used for lunch and a daily nap. Beyond this, teachers could use the remaining hours in a variety of ways including literacy instruction, math instruction, structured or unstructured play, etc. Aside from the substantial differences with respect to intensity, full- and half-day classrooms were similar in many respects. All WPS pre-k classrooms were led by teachers with a bachelor's degree, they maintained the same teacher-to-child ratios, and they used the same curriculum (Literacy Express).

This paper presents findings from the first year of FDPK, and is focused on children's school readiness outcomes, as measured both at the end of the preschool year and at the beginning of kindergarten. Below we describe the research design for the evaluation study, analysis models, and sensitivity checks for RCT estimates.

## **Methods**

### **Research Design**

The current study of FDPK employs a randomized block design (within first choice of school site) in which eligible families who completed an application were randomly assigned to offers of full- and half-day classrooms. In this case, the block randomized design was ideal because it allowed the study team to accommodate families' preferences for school sites, while also reducing the likelihood for chance imbalances across groups. Because a small number of families did not take-up their lottery assignments into full- or half-day classrooms, we estimate the intent-to-treat (ITT) and complier-average-treatment-effects (CATE) as the causal estimands of interest.

### **Sample**

All children who reside within the district and are age four by the first of October are eligible to participate in the FDPK program. In order for families to enroll in WPS generally, they must complete a required preschool application that includes health certifications and the child birth certificate. Families were included in the current study if they expressed interest in full-day preschool on their application, they completed the consent process, and their child(ren) had no known special education needs that prevented them from being served within a full-day classroom (e.g., if special equipment was required, a six-hour day is inappropriate).

The sample includes a total of 226 children with 114 children randomly assigned to offers for FDPK and 112 assigned to offers for the half-day program. Table 1 provides descriptive information about the study sample. Overall, the sample is largely Hispanic (74 percent) and low-income (61 percent qualified for free-lunch, and 13 percent qualified for reduced-price lunch). The General Preschool Application (more on this data source below) asks the child's primary caregiver a series of questions about the child's family history. For instance, 37 percent reported having received some education beyond high school, and 51 percent indicated that their home language is English. About 20 percent responded 'yes' to the question, "Has an immediate family member [of the child] received Special Education services?". About 20 percent of caregivers also indicated that the enrolled student has low language development, and 40 percent indicated low social development for the child.<sup>2</sup> The average age of children enrolled in the pilot study was 4.39 years, and about half the children are male.

### **Data Collection**

To examine the impact of FDPK on children's outcomes, the study team assessed children's receptive vocabulary skills and administered an intensive developmental screener that identifies children who may need special education services. These assessments were conducted within the first month of fall 2016 (baseline) and again in the last month of spring 2017 (end of pre-k year). Crucially, all study children were administered the same assessments, regardless of whether or not they enrolled in WPS pre-k. If a child was enrolled in WPS pre-k (half- and full-day programs), the study team administered the receptive vocabulary assessment and developmental screener during regular pre-k hours. If a child was not enrolled in WPS, the study team met with the family directly to administer measures outside of the classroom setting.

In addition to outcome measures collected by the study team, the evaluation also includes assessments administered directly by the school district. However, because these measures are available only for the subset of children enrolled in WPS, we treat these study results as exploratory.

### **Measures of School Readiness**

**Primary outcomes.** Children's receptive vocabulary was measured by the Peabody Picture Vocabulary Test, 4<sup>th</sup> Edition (PPVT-4) (L. M. Dunn & Dunn, 2007). The PPVT-IV is a 228-item test in standard English administered by having children point to one of four pictures that best corresponds to a spoken word. The PPVT-4 Scale is norm-referenced and is widely used as a measure of children and adult's receptive, or heard, vocabulary. The PPVT-IV has strong psychometric properties with evidence for high reliability and validity (L. Dunn & Dunn, 2013) (Dunn & Dunn, 2013).

The Early Screening Inventory- Revised (ESI-R) is a one-on-one, 20-minute developmental screening tool that is appropriate for children from 3 years 5 months to 5-years 11 months (Meisels, Marsden, Wiske, & Henderson, 1997). The measure evaluates children's developmental abilities in three domains of school readiness. To assess visual-motor/adaptive reasoning, the child is asked to replicate patterns with blocks and copy with a drawing. To assess cognition and language, the child is given four tasks that allow her to demonstrate ability to comprehend language, express ideas, and reason and count. To assess gross motor skills, the child is asked to jump, hop, and other physical coordination tasks. Results from the ESI-R indicate the instrument is both reliable and valid (Meisels, Henderson, Liaw, Browning, & Ten Have, 1993; Moodie et al., 2014). A Spanish language version of the screener is also available, and we administer the ESI-R in the child's primary language.

**Exploratory outcomes.** As discussed above, in addition to the data collected directly by our research team, we also considered two outcomes collected by WPS. First during the fall and spring of the pre-k year, all teachers assessed children using Teaching Strategies GOLD (TS GOLD) a widely-used, observation-based authentic assessment (Heroman et al., 2010). Teachers observe children's skills during typical classroom sessions and evaluate them across up to nine broad areas of development (e.g., literacy, mathematics, language, social-emotional, cognitive, and physical). TS GOLD has been used in other studies tracking the association between preschool intensity and child outcomes (Reynolds et al., 2014). Although teacher-reported, the measure is correlated with direct assessments and has shown strong reliability and validity (Teaching Strategies, 2011, 2013). Further, recent research indicates the assessment functions well with children whose home language is not English (Kim, Lambert, & Burts, 2013).

The second district-administered measure we use in this study is the Dynamic Indicators of Basic Early Literacy Skills (DIBELS) (Good & Kaminski, 2002). The DIBELS assesses children's early literacy skills in the areas of phonemic awareness, phonics, reading comprehensive, fluency, and vocabulary. It is designed to help identify children who may experience difficulty acquiring basic early literacy skills. The measure shows adequate validity and reliability and has been adopted by a number of states to assess children's school readiness (Good et al., 2004). WPS administers the DIBELS to children during the fall and spring of the kindergarten year and provided us with two subdomain scores—First Sound Fluency and Letter Naming Fluency—as well as an overall composite score. The current study presents RCT results on the TS GOLD at the end of the pre-k year, and the DIBELS in the fall of the kindergarten year.

### **Baseline Measures**

A unique strength of the current study is that we have access to an unusually rich set of baseline covariates, including measures of child, home, and family characteristics as well as baseline assessments for primary outcomes. All applicants to Westminster complete a General Preschool Application, which included questions about children's race/ethnicity, gender, birthdate and age, free/reduced-price lunch program eligibility, primary language, and the primary language spoken in the home. In addition, the child's parent or primary guardian indicated their educational background, and whether there is a history of family drug/alcohol abuse, special needs, frequent school moves, housing difficulty, domestic abuse, social services involvement, and extreme child medical events occurring within the family. Parents also indicated whether they were concerned about the child's low language and/or social development. Finally, our research team administered measures of the PPVT and the ESI-R in early fall of the pre-k year and teachers also assessed children using the TS GOLD during the same period. These baseline covariates greatly enhance our ability to assess covariate balance across groups and improve statistical precision to detect effects.

### **Analysis Model**

In Equation (1), we present the statistical model used to estimate causal effects of full-day pre-k on student outcomes of interest:

$$Y_{ijt=1} = \beta_0 + \beta_1(T_{ij}) + (\mathbf{X}_{ij(t=0)})\boldsymbol{\beta} + \alpha_j + \varepsilon_{ijt}. \quad (1)$$

$Y_{ijt+1}$  represents an outcome for student  $i$  in school site  $j$  at the end of the second semester of the pre-k year ( $t=1$ ). The variable  $T_{ij}$  is the dummy variable coded to 1 if the child was randomly assigned to receive an offer of full-day pre-k (treatment) and 0 if offered a half-day spot (control). Our analysis models also include a series of school site fixed effects  $\alpha_j$ , as well as a vector of time-invariant ( $\mathbf{X}_{ij}$ ) and time-varying controls ( $\mathbf{X}_{ijt=0}$ ) that were taken at baseline (see Table 1 for a list



of covariates). In this model,  $\beta_1$  captures the intent-to-treat effect of interest—that is, conditional on covariates in the model,  $\beta_1$  is the average effect on  $Y$  due to randomly assigned offers of full- or half-day pre-k. We also used two-stage least square approaches to estimate the complier average treatment effects—that is, it is the effect of *attending* full-day pre-k on children’s school readiness skills.

### **Validity Checks for Experimental Design**

Because randomized experiments in field settings are rarely (if ever) implemented perfectly, we conducted a series of diagnostic probes to assess the extent to which validity threats occurred. Below, we discuss the results of our diagnostic checks for covariate balance, treatment non-compliance, and for missing data on outcomes.

**Covariate balance.** Even with random assignment, it is possible the full- and half-day groups differ based on chance alone. We tested whether there were differences in groups at baseline by running a series of regressions in which each baseline covariate was regressed on indicators for whether the family was offered full-day pre-k and site fixed effects. The dependent variables in these regressions included all baseline covariates discussed above. Each row of Table 2 presents the results from a separate regression with the same specification but a different dependent variable. Results in a standardized metric are also presented visually in Figure 1

**Error!**  
**Reference source not found..**

The full- and half-day groups were similar at baseline. For instance, the groups are very well matched with respect to racial composition, age, gender, home language, and parental education. As expected by chance, only one of the twenty-one group differences considered is statistically significant. The one statistically significant coefficient is for the percentage of families reporting “frequent moves” (5 percent of treatment children, and 17 percent of control children,

and an associated p-value of 0.003). Furthermore, there does not appear to be any systematic patterns of advantage or disadvantage between the groups. The full-day group has somewhat higher percentages of some characteristics that are historically associated with lower test score performance (e.g., children in the full-day group are approximately 11 percentage points more likely to be eligible for free lunch). However, in other cases, half-day families exhibit slightly higher means (e.g., 75.8 percent of control families are Hispanic, while 72.0 percent of treatment families are Hispanic). Overall, the balance tests suggest little evidence that there were systematic differences between groups at baseline. Further, sensitivity analyses indicate that the magnitude of our experimental estimates is robust to the inclusion of baseline covariates in the model.

**Treatment non-compliance.** Families are randomized to *offers* of full or half-day pre-k slots in WPS. Some may choose to not take up their offers. In particular, families assigned to a half-day WPS slot may opt out of WPS pre-k and could possibly enroll their child in a different full day setting. Such treatment non-compliance leads to a discrepancy between assigned and observed treatment status. Across both conditions, 74 percent of the study sample participated in the pre-k classroom to which they were assigned. Among those who were randomly assigned to full-day pre-k, 86 percent attended the full-day program in WPS. Among those assigned to the half-day group, 62 percent participated in half-day classes in WPS. This differential take-up rate across groups was expected given that all study families had initially indicated interest in a full-day slot. A small portion of study participants experienced crossover. Specifically, two percent of families assigned to full-day pre-k switched to the half-day program in WPS, and nine percent of families who were initially assigned to half-day pre-k enrolled in the WPS full-day program.

To address threats to internal validity from this non-compliance we estimate ITT effects that isolate the causal impact of receiving an *offer* to participate in the full-day program, as well as CATE effects, which estimate the impact of the program for compliers.

**Missing data and attrition.** Because families were required to complete the General Preschool Application to enroll in FKPK, there is very little missing data on baseline covariates (see Table 1 for missing data rates). For the few cases in which we did not have baseline covariate information, we included controls for missing data in the analysis models.

Table 1 shows that 12 percent of the study sample is missing data on the main student outcomes of interest, the spring PPVT and ESI-R because we were unable to reach some families to complete the assessments. Table 3 summarizes baseline characteristics of half- and full-day children for the full sample, and for the remaining sample without missing information. It shows that the remaining sample is similar to the complete study sample, and that there is no evidence of differential attrition across the groups.

As expected, there is more missing data on assessments that were administered by the school district rather than by our study team. WPS administers TS GOLD to children during the fall and spring of the pre-k year, and DIBELS during the fall of the kindergarten year. Because the assessments are given only to participants who enrolled in public pre-k and kindergarten, we lack these outcomes for children who did not participate in WPS pre-k and/or kindergarten. Altogether, we lack TS GOLD data for approximately 20 percent of the study sample, and DIBELS data for 38 percent of the sample.

Tables A1 and A2 in the Appendix A compare baseline information for the full RCT sample, and the subset of children who had TS GOLD scores at the end of the pre-k year and DIBELS scores in the fall of the kindergarten year, respectively. Reassuringly, we find that

although there is substantial missing data across both groups, we do not observe systematic differences in either full- or half-day group, suggesting that the kinds of participants missing data may generally not systematically differ from the full sample. Nevertheless, given the high levels of missingness, we treat these outcomes as exploratory. In specification checks, described further below, we also assess whether impacts on these outcomes are robust to conservative assumptions about the nature of the missing data.

## **Results**

Table 4 presents our estimates of the ITT and CATE for our primary outcomes, PPVT and ESI-R. Columns labeled M1 show the standardized mean differences in effects with only school site fixed effects included in the model; the M2 columns show impact estimates with school site fixed effects and controls for student and family demographic factors; and the M3 columns present results from our preferred model, which includes school site fixed effects, controls for demographic factors, as well as baseline pretest scores. Across all three models, treatment effect estimates are generally stable, while the proportion of variance explained increases substantially across models.

Table 4 shows that the impact of offering full-day pre-k resulted in an increase of 0.267 standard deviations on the PPVT-4 (upper-left panel of Table 4, Model 3). The impact of actually attending full-day pre-k improved children's receptive vocabulary scores by 0.398 standard deviations (lower-left panel of Table 4, Model 3). For ESI-R, all estimates across models and estimands are positive and between 0.40 and 0.92 standard deviations. The ITT effect is 0.054 standard deviations and the CATE result is 0.080 standard deviations (upper- and lower-right panels of Table 4, Model 3). Neither are statistically significant.

Table 5 contains ITT and CATE results for TS GOLD and DIBELS, the outcome measures collected only for those children enrolled in WPS. For the sake of parsimony, we only present results in Table 5 from our preferred Model 3. Table 5 results suggest that children randomly assigned to an offer of full-day pre-k were rated more highly on the TS GOLD than their peers in half-day programs. Treatment effects are positive for all six domains assessed, and statistically significant for four of the six domains (cognition, literacy, math, and physical development). The largest effects were for literacy (ITT = 0.390 standard deviations; CATE = 0.489 standard deviations), followed by math (ITT = 0.197 standard deviations, CATE = 0.247 standard deviations), cognition (ITT = 0.191 standard deviations, CATE = 0.239 standard deviations), and physical development (ITT = 0.179 standard deviations, CATE = 0.224 standard deviations). TS GOLD scores on language and socio-emotional skills are substantively large and positive but do not differ significantly across groups.

Table 5 also indicates that by the fall of kindergarten children randomly assigned to an offer of full day pre-k outperformed their peers on the DIBELS. The ITT effect for the overall composite score of the DIBELS was 0.345 standard deviations, and for the CATE, it was 0.392 standard deviations. We also see positive effects on the two provided DIBELS subdomains—First Sound Fluency (ITT effect= 0.349) and Letter Naming Fluency (ITT effect= 0.330). With the exception of Letter Naming Fluency, all the DIBELS effects were statistically significant at the 10 percent alpha level.

### **Robustness Checks for Missing Outcome Data**

Despite the fact that we make every effort to assess all study children not enrolled in WPS, we do not observe outcomes for all study children at the end of preschool or the start of kindergarten due to the natural mobility that occurs in any district during and after preschool (see

Table 1). One could be concerned that the missingness is systematic and could bias our results. We would be particularly concerned in a scenario where high-scoring control students (or low-performing treatment students) were more likely to have missing data, as these patterns would bias our estimates upward so that they appear larger than they actually are.

As a robustness check for our estimated effects, we make assumptions about the missing outcome scores that would work strongly *against* our findings: Within each school, we assume that every missing control group child would have performed on these assessments at the average level of the apparently higher-scoring treatment group. Likewise, we assume that every missing treatment group child would have performed at the average level of the control group. These strong assumptions correspond to the upwards bias scenario described above. Recall that we do see evidence that the observable characteristics of the pre- and post-missing data samples are systematically different from one another, so this thought experiment may be somewhat overly-punitive. Nevertheless, we can look to see if the direction and magnitude of estimated effects under these assumptions remain positive and meaningfully large (if not still statistically significant).

When we make these assumptions, we find the pattern of the results generally persists. We rerun the analyses presented in Table 4 and Table 5 now with the imputed outcome data and present updated results in Table 6. Note that in Table 6 there are now N=226 children in every model because all study participants now have a value for the outcomes, imputed or otherwise. In Table 6 we see that the estimated effects are generally smaller in magnitude, but most remain positive and substantively large. Statistical significance should be interpreted with caution when analyzing imputed outcome data, however ten of the estimated effects continue to be statistically significant despite moving closer to zero. This suggests that the direction and magnitude of our findings are

insensitive to the missing data that is both present in this study and endemic to all longitudinal early childhood research designs.

## **Discussion**

To complement the large body of research examining *whether* preschool leads to benefits for children, evidence is needed on the *conditions* under which preschool is most effective. The current study provides the first rigorous evidence on the effects of full-day preschool on young children's school readiness. Unlike the majority of the existing research on the intensity of early childhood interventions that reports regression adjusted associations between program exposure and child outcomes, the current study leverages a school-based lottery to conduct an RCT, thus isolating the causal relationship between full-day programs and young children's early development.

The results indicate that the offer of full-day pre-k positively impacted young children's school readiness skills. In particular, children offered full-day pre-k scored a quarter of a standard deviation higher on the PPVT, a widely-used measure of receptive vocabulary, than peers offered half-day pre-k. These impacts are sizable. To put them in perspective, we compare them to rigorous findings from studies examining the overall impact of ECE interventions, rather than the specific impact of program intensity. Experimental evidence on the impacts of Head Start indicated that three-year-olds randomly assigned to a Head Start slot scored 0.13 standard deviations higher than those in the control group on the PPVT (Puma, Bell, Cook, Heid, & Lopez, 2005). Among four-year-olds, random assignment to a Head Start offer was not significantly related with this measure. Wong et al. (2008) used regression discontinuity methods to estimate the impact of five state pre-kindergarten programs on the PPVT. They found that effects sizes ranged from a statistically insignificant -0.13 in Michigan to a statistically significant 0.36 in New Jersey. Only two of the

five states considered showed statistically significant positive impacts on this outcome. In a recent study expanding this work to eight state pre-kindergarten programs (Barnett et al., 2018), the average effect sizes of pre-kindergarten on the PPVT was 0.24, though only three of eight states (New Jersey, Michigan and Oklahoma) showed statistically significant impacts in the authors' preferred model, and results were sensitive to model fit. The effects observed in the current study are thus larger than what is often observed in studies of overall ECE impacts, and roughly the same size as those seen for some of the most successful state pre-kindergarten programs. These findings are encouraging, especially given the importance of unconstrained skills, and particularly early vocabulary, for children's reading at third grade and longer-term literacy success (Snow & Matthews, 2016).

Our results indicate positive but statistically insignificant effects at the end of the preschool year on the ESI-R, a developmental screener used to identify children who may need special education services. We find positive outcomes with respect to the TS GOLD, a widely-used, teacher-reported observational tool, which captures development across a broader range of developmental domains. Here too effects were encouraging. We found statistically significant and sizable impacts on cognition, literacy, math, and physical skills, ranging in effects sizes from 0.19 to 0.39. The coefficients for the other two outcomes considered (language and socio-emotional development) were both about 0.10 and were not statistically significant. A recent nonrandomized study exploring the impacts of full-day preschool in the context of Chicago's Child-Parent Centers also showed that children in full-day classrooms outperformed their peers on four of six TS GOLD outcomes, though curiously they found statistically significant outcomes on language and socio-emotional development but not literacy or cognition (Reynolds et al., 2014).



We interpret the TS GOLD findings with caution for two reasons. First, because the measure was collected as part of “business-as-usual” practice for WPS pre-k, it is unavailable for the non-random sample of children who ultimately did not attend pre-k in WPS despite receiving an offer to do so. Findings in the current analysis are robust to relatively conservative assumptions about the values of this missing data. A second concern about the TS GOLD data is that it is reported by teachers who are aware of children’s full versus half-day status, which may introduce bias. Several factors may work against this possibility. First, the formal TS GOLD trainings teachers take are designed specifically to increase accuracy and reduce rating bias. In addition, WPS teachers have been routinely administering TS GOLD in their classrooms for at least six years prior to the study and therefore it is very much standard practice and not specific to this study. Ultimately we cannot directly address possible bias in teacher-reported assessments, however existing research does suggest that teachers’ ratings of children are strongly correlated with direct assessments.

Finally, effect sizes for the DIBELS, a direct literacy assessment administered by WPS in the fall of the kindergarten year, were substantial ( $ES=0.34$ ), though only marginally statistically significant given the smaller sample size. These results closely align with C. R. Gibbs (2014) whose lottery-based analysis of full-day kindergarten shows ITT effects of approximately a third of a standard deviation. Again, these findings are encouraging given the association between early literacy, as measured by this assessment, and children’s development of reading skills throughout elementary school (Burke, Hagan-Burke, Kwok, & Parker, 2009; Rouse & Fantuzzo, 2006). However, as with the TS GOLD, caution is warranted, given the relatively high rates of missingness on this WPS-only outcome.

Taken together, the effects documented in the current paper, which were systematically positive, and in most cases also statistically significant, provide the most rigorous evidence to date on the impacts of an extended pre-kindergarten day for young children's school readiness skills. These findings are important, especially in light of recent calls for more rigorous evidence on the impacts of specific aspects of ECE in fostering children's learning gains (Weiland, 2018).

At the same time, the current study, which focuses on the immediate impacts of full-day pre-k on child outcomes within one Colorado district, leaves many important questions unanswered, and four of these warrant particular consideration: (1) To what extent will the benefits observed at the end of the pre-k year and at the beginning of kindergarten be maintained as children proceed through the early grades and beyond?; (2) To what extent do the experimental findings documented in the current study—which focused on a predominantly Hispanic, predominantly low-income sample—generalize to other contexts?; (3) If full-day classrooms are effective, what specific practices and experiences are driving the benefits and; (4) How do the costs and benefits associated with full-day pre-k compare to the returns from other approaches to ensuring ECE quality? Each of these questions are critical towards understanding the promise of expanded access to pre-k as a strategy for supporting young children's development.

In ongoing work we are beginning to answer these questions by tracking the children in the current study as they proceed through the first four years of elementary school. In addition to the current cohort of children, we are tracking two additional cohorts of WPS children, who will also be randomly assigned to an offer of full- or half-day pre-k. For these cohorts we are collecting unusually rich data that will allow us to descriptively unpack the *mechanisms* driving the benefits of a full-day pre-k experience. In particular, through multiple classroom observations throughout the year, we are collecting detailed information about the time-use in full and half day classrooms,

as well as the quality of teacher-child interactions as measured by the Classroom Assessment Scoring System, CLASS (Pianta, La Paro, & Hamre, 2008). We also supplement these observational measures with detailed parent surveys which allow us to explore the potential impacts of full-day programs on parental employment and family well-being. These surveys will also provide us with detailed information about the counterfactual condition, highlighting how young children in half-day programs spend the out-of-school portions of their day. Finally, we are in the process of replicating our study in the context of a second school district located in southern California, which also serves a predominantly Hispanic population, but is targeted and serves a substantially more at-risk sample.

Our goal is that through this deep-dive into the impacts of one particular feature of ECE, program quality, and through similar undertakings about other potentially central ECE features such as curricula, professional development, etc., we will begin to provide policy-makers with the kind of evidence necessary to make smart decisions not about whether or not to offer ECE programs but about *how* to design policies that yield meaningful and sustained impacts.

## Tables

**Table 1. Study Sample Descriptive Statistics, Relative to Full DPS District**

	Pre-K Study Sample			
	Mean	SD	N	% Missing
<i>Baseline Demographics</i>				
% Male	0.49	0.50	226	0.0
% Red-Lunch Elig	0.13	0.34	226	0.0
% Free-Lunch Elig	0.61	0.49	226	0.0
% Unknown Lunch Status	0.22	0.41	226	0.0
% Parent Ed > HS	0.37	0.48	226	0.0
% with Teen Parents	0.02	0.13	226	0.0
% Hispanic	0.74	0.44	226	0.0
% Black	0.00	0.07	226	0.0
% Home Lang. English	0.51	0.50	214	5.3
% with Fam Hist of Drug/Alc Abuse	0.01	0.11	226	0.0
% with Hist of Abuse	0.01	0.09	226	0.0
% with Fam History of Special Needs	0.17	0.37	226	0.0
% with Fam Hist of Frequent Moves	0.11	0.31	226	0.0
% with Child Extreme Medical Event	0.02	0.15	226	0.0
% with Housing Difficulty	0.04	0.19	226	0.0
% with Low Language Development	0.23	0.42	226	0.0
% with Low Social Development	0.37	0.48	226	0.0
% Receiving Social Services	0.03	0.16	226	0.0
Age	4.36	0.30	223	1.3
<i>Assessment Variables</i>				
PPVT PK Fall Std. Score	92.1	20.1	200	11.5
PPVT PK Spring Std. Score	96.2	19.1	200	11.5
ESI-R PK Fall Total Score	19.0	6.7	215	4.9
ESI-R PK Spring Total Score	20.6	5.1	202	10.6
TS GOLD PK Fall Cognitive Score	568.6	55.9	179	20.8
TS GOLD PK Fall Language Score	560.3	52.2	179	20.8
TS GOLD PK Fall Literacy Score	563.5	44.2	179	20.8
TS GOLD PK Fall Math Score	567.3	46.0	178	21.2
TS GOLD PK Fall Phys. Dev. Score	564.8	44.6	179	20.8
TS GOLD PK Fall Soc. Emot. Score	576.8	49.4	179	20.8
TS GOLD PK Spring Cognitive Score	676.2	63.5	182	19.5
TS GOLD PK Spring Language Score	656.5	63.1	182	19.5
TS GOLD PK Spring Literacy Score	653.2	54.5	182	19.5
TS GOLD PK Spring Math Score	656.4	54.7	182	19.5
TS GOLD PK Spring Phys. Dev. Score	647.8	57.3	182	19.5
TS GOLD PK Spring Soc. Emot. Score	666.8	59.7	182	19.5
DIBELS K Fall Composite Score	22.3	20.4	139	38.5

*FN: Descriptive statistics for the study sample. Note that 226 study children were randomized. Demographic and family history questions come from the general application for WPS preschool, which is required for all WPS preschool applicants. To see exact wording of these questions, see the Footnote of Table 2.*

**Table 2. Baseline Covariate Balance, Expressed in Original Metrics and Standardized Cohen's D**

Pre-Treatment Covariate	Treatment	Control	Raw	Cohen's D	T-Statistic	P-Value	N	95% CI	95% CI
	Mean	Mean	Difference					Low	High
% White	11.4%	11.6%	-0.1%	-0.004	0.032	0.974 --	226	-0.086	0.083
% Black	0.8%	0.0%	0.8%	0.122	0.908	0.365 --	226	-0.009	0.026
% Hispanic	72.0%	75.8%	-3.8%	-0.086	0.639	0.523 --	226	-0.153	0.078
% Home Lang. Not English	48.8%	48.4%	0.4%	0.008	0.058	0.954 --	214	-0.130	0.138
% Parent Ed > HS	36.4%	38.0%	-1.6%	-0.033	0.244	0.807 --	226	-0.145	0.113
% with Teen Parents	1.8%	1.8%	0.0%	0.003	0.019	0.985 --	226	-0.034	0.035
% Free-Lunch Elig	66.2%	55.0%	11.2%	0.229	1.724	0.086 --	226	-0.015	0.240
% Unknown Lunch Status	21.8%	21.5%	0.3%	0.007	0.056	0.955 --	226	-0.103	0.109
% Red-Lunch Elig	11.3%	14.4%	-3.2%	-0.095	0.703	0.483 --	226	-0.120	0.057
% Male	48.0%	49.3%	-1.3%	-0.025	0.186	0.852 --	226	-0.145	0.120
Child's Age (in yrs)	4.34	4.38	-0.04	-0.124	0.923	0.357 --	223	-0.118	0.043
% with Hist of Abuse	1.9%	-0.2%	2.1%	0.224	1.664	0.097 --	226	-0.004	0.046
% with Fam Hist of Drug/Alc Abuse	0.9%	1.7%	-0.8%	-0.070	0.524	0.601 --	226	-0.038	0.022
% with Fam History of Special Needs	17.9%	15.8%	2.1%	0.056	0.414	0.679 --	226	-0.078	0.120
% with Child Extreme Medical Event	0.9%	3.6%	-2.7%	-0.182	1.351	0.178 --	226	-0.066	0.012
% with Fam Hist of Frequent Moves	4.6%	16.7%	-12.1%	-0.391	2.970	0.003 **	226	-0.200	-0.041
% with Housing Difficulty	4.7%	2.3%	2.4%	0.131	0.969	0.333 --	226	-0.025	0.073
% with Low Language Development	24.3%	21.7%	2.6%	0.062	0.459	0.646 --	226	-0.086	0.138
% with Low Social Development	38.1%	35.4%	2.7%	0.056	0.430	0.668 --	226	-0.097	0.152
% Received(ing) Social Services	2.6%	2.7%	0.0%	-0.001	0.011	0.991 --	226	-0.043	0.043
Fall PPVT Std. Score	0.027	-0.034	0.061	0.061	0.461	0.645 --	215	-0.198	0.319
Fall ESI-R Total Score	0.084	-0.093	0.178	0.178	1.321	0.188 --	215	-0.086	0.441

FN: Demographic and family history questions come from the general application for WPS preschool, which is required for all WPS preschool applicants. Parents are asked to answer yes or no to the following questions: Q1: Is there or has there been drug or alcohol abuse in the home where the child lives? Q2: Is there or has there been an abusive adult residing in the home? Q3: Has an immediate family member received Special Education services? Q4: Has your child moved frequently? Q5: Has your child had any serious illness or medical diagnosis? Q6: Is your child in need of language development including, but not limited to, the ability to speak English? Q7: Does your child have problems with social situations? Q8: Is or has your child been receiving services from the State Department of Social Services related to neglect or abuse? Finally, we coded a family as having a housing difficulty if they answered yes to a series of questions asking if the family is in a temporary housing situation due to loss of housing or economic hardship.

**Table 3. Baseline Covariate Balance Comparison: Full Sample vs. Sample with End-of-Pre-Kindergarten Outcomes**

Pre-Treatment Covariate	Treatment Group				Control Group			
	Full		Non-		Full		Non-	
	Mean	(N)	Mean	(N)	Mean	(N)	Mean	(N)
% White	11.4%	(114)	12.3%	(107)	11.6%	(112)	9.2%	(96)
% Black	0.8%	(114)	0.9%	(107)	0.0%	(112)	0.0%	(96)
% Hispanic	72.0%	(114)	72.3%	(107)	75.8%	(112)	77.8%	(96)
% Home Lang. Not English	48.8%	(112)	47.8%	(105)	48.4%	(102)	49.8%	(88)
% Parent Ed > HS	36.4%	(114)	35.3%	(107)	38.0%	(112)	42.9%	(96)
% with Teen Parents	1.8%	(114)	2.0%	(107)	1.8%	(112)	2.0%	(96)
% Free-Lunch Elig	66.2%	(114)	65.1%	(107)	55.0%	(112)	57.6%	(96)
% Unknown Lunch Status	21.8%	(114)	21.8%	(107)	21.5%	(112)	17.4%	(96)
% Red-Lunch Elig	11.3%	(114)	12.2%	(107)	14.4%	(112)	15.6%	(96)
% Male	48.0%	(114)	50.5%	(107)	49.3%	(112)	48.9%	(96)
Child's Age (in yrs)	4.34	(114)	4.35	(107)	4.38	(109)	4.36	(93)
% with Hist of Abuse	1.9%	(114)	2.0%	(107)	-0.2%	(112)	-0.1%	(96)
% with Fam Hist of Drug/Alc Abuse	0.9%	(114)	1.0%	(107)	1.7%	(112)	2.0%	(96)
% with Fam History of Special Needs	17.9%	(114)	18.0%	(107)	15.8%	(112)	17.5%	(96)
% with Child Extreme Medical Event	0.9%	(114)	1.0%	(107)	3.6%	(112)	4.1%	(96)
% with Fam Hist of Frequent Moves	4.6%	(114)	5.3%	(107)	16.7%	(112)	18.1%	(96)
% with Housing Difficulty	4.7%	(114)	5.0%	(107)	2.3%	(112)	2.8%	(96)
% with Low Language Development	24.3%	(114)	24.1%	(107)	21.7%	(112)	23.1%	(96)
% with Low Social Development	38.1%	(114)	39.0%	(107)	35.4%	(112)	34.7%	(96)
% Received(ing) Social Services	2.6%	(114)	2.9%	(107)	2.7%	(112)	3.1%	(96)
Fall PPVT Std. Score	0.027	(111)	0.015	(107)	-0.034	(104)	-0.093	(92)
Fall ESI-R Total Score	0.084	(111)	0.103	(107)	-0.093	(104)	-0.050	(92)

*Note: "Full" sample refers all study participants originally assigned to each treatment group. The "Non-Attrition" sample refers to study participants who are observed with PPVT test scores in the spring of pre-k.*

**Table 4. Primary Outcomes (End of Pre-K): Causal Effects of Full-Day Pre-kindergarten (ITT vs. CATE)**

<i>Intent-to-Treat Analysis</i>						
	<i>PPVT End of Pre-K</i>			<i>ESI-R End of Pre-K</i>		
	<i>(M1)</i>	<i>(M2)</i>	<i>(M3)</i>	<i>(M1)</i>	<i>(M2)</i>	<i>(M3)</i>
<b>Assigned to Full</b>	0.248 *	0.261 *	0.267 **	0.040	0.056	0.054
	(0.136)	(0.131)	(0.101)	(0.143)	(0.147)	(0.148)
<b>Constant</b>	0.004	0.117	0.160 *	0.013	0.074	0.166
	(0.099)	(0.098)	(0.079)	(0.103)	(0.110)	(0.116)
<b>R<sup>2</sup></b>	0.055	0.394	0.716	0.026	0.295	0.431
<b>Adj. R<sup>2</sup></b>	0.020	0.273	0.637	-0.009	0.151	0.271
<b>N</b>	200	200	200	202	202	202
<i>Complier Average Treatment Effects (Two-Stage Least Squares)</i>						
	<i>PPVT End of Pre-K</i>			<i>ESI-R End of Pre-K</i>		
	<i>(M1)</i>	<i>(M2)</i>	<i>(M3)</i>	<i>(M1)</i>	<i>(M2)</i>	<i>(M3)</i>
<b>Attended Full-Day</b>	0.341 *	0.355 *	0.398 **	0.074	0.092	0.080
	(0.172)	(0.179)	(0.153)	(0.178)	(0.197)	(0.221)
<b>Constant</b>	-0.044	0.064	0.063	-0.005	0.053	0.146
	(0.113)	(0.118)	(0.109)	(0.117)	(0.131)	(0.158)
<b>N</b>	200	200	200	202	202	202

*Note: M1 includes first-choice school (i.e., block) fixed effects only. In M2, we add student-level control variables (for the sake of power) for gender, race/ethnicity, free/reduced price lunch eligibility, and age at the start of pre-k. To control for any possible family context differences we also include parental education, whether or not child had a teen parent, whether the home language is English, as well as indicators for (pre-treatment reports of) history of family drug/alcohol abuse, family special needs, frequent moving, physical abuse, interactions with social services, difficulties obtaining secure housing, and perceived child underdeveloped language or social skills prior to pre-k. In M3, we add start of schoolyear pre-scores on all assessments (and, where relevant, the language and format in which these assessments were taken). We include missingness dummies in cases where respondents have missing pre-treatment covariates.*

**Table 5. Exploratory Outcomes: Causal Effects of Full-Day Pre-kindergarten (ITT vs. CATE), Full Model (3) Only**

Intent-to-Treat Analysis									
	End of Pre-K TS GOLD						Fall Kindergarten DIBELS		
	Cognition	Language	Literacy	Math	Physical	Socio-Emotional	Overall Composite	1st Sound Fluency	Ltr Naming Fluency
Assigned to Full	0.191 *	0.092	0.390 ***	0.197 *	0.179 *	0.100	0.345 +	0.349 +	0.290
	(0.082)	(0.086)	(0.086)	(0.080)	(0.085)	(0.073)	(0.178)	(0.203)	(0.182)
Constant	0.106	0.113	0.152 *	0.088	-0.001	0.033	0.237 +	0.190	0.239 +
	(0.065)	(0.069)	(0.068)	(0.064)	(0.068)	(0.058)	(0.137)	(0.157)	(0.141)
R^2	0.799	0.774	0.827	0.831	0.778	0.832	0.611	0.540	0.568
Adj. R^2	0.734	0.702	0.771	0.777	0.707	0.778	0.441	0.338	0.379
N	182	182	182	182	182	182	139	139	139

*Complier Average Treatment Effects (Two-Stage Least Squares)*

	End of Pre-K TS GOLD						Fall Kindergarten DIBELS		
	<i>Cognition</i>	<i>Language</i>	<i>Literacy</i>	<i>Math</i>	<i>Physical</i>	<i>Socio-Emotional</i>	<i>Overall Composite</i>	<i>1st Sound Fluency</i>	<i>Ltr Naming Fluency</i>
<b>Attended Full-Day</b>	0.239 *	0.115	0.489 ***	0.247 *	0.224 *	0.125	0.392 +	0.397 +	0.330
	(0.102)	(0.107)	(0.104)	(0.098)	(0.107)	(0.091)	(0.202)	(0.230)	(0.207)
<b>Constant</b>	0.067	0.094	0.073	0.048	-0.037	0.013	0.204	0.155	0.211
	(0.077)	(0.081)	(0.079)	(0.074)	(0.081)	(0.069)	(0.152)	(0.173)	(0.155)
<b>N</b>	182	182	182	182	182	182	139	139	139

*Note: Results are reported for M3 only, which includes first-choice school (i.e., block) fixed effects, student-level control variables, and start of schoolyear pre-scores on all assessments (and, where relevant, the language and format in which these assessments were taken). The vector of student-level pre-treatment covariates includes gender, race/ethnicity, free/reduced price lunch eligibility, and age at the start of pre-k. To control for any possible family context differences we also include parental education, whether or not child had a teen parent, whether the home language is English, as well as indicators for (pre-treatment reports of) history of family drug/alcohol abuse, family special needs, frequent moving, physical abuse, interactions with social services, difficulties obtaining secure housing, and perceived child underdeveloped language or social skills prior to pre-k. We include missingness dummies in cases where respondents have missing pre-treatment covariates.*



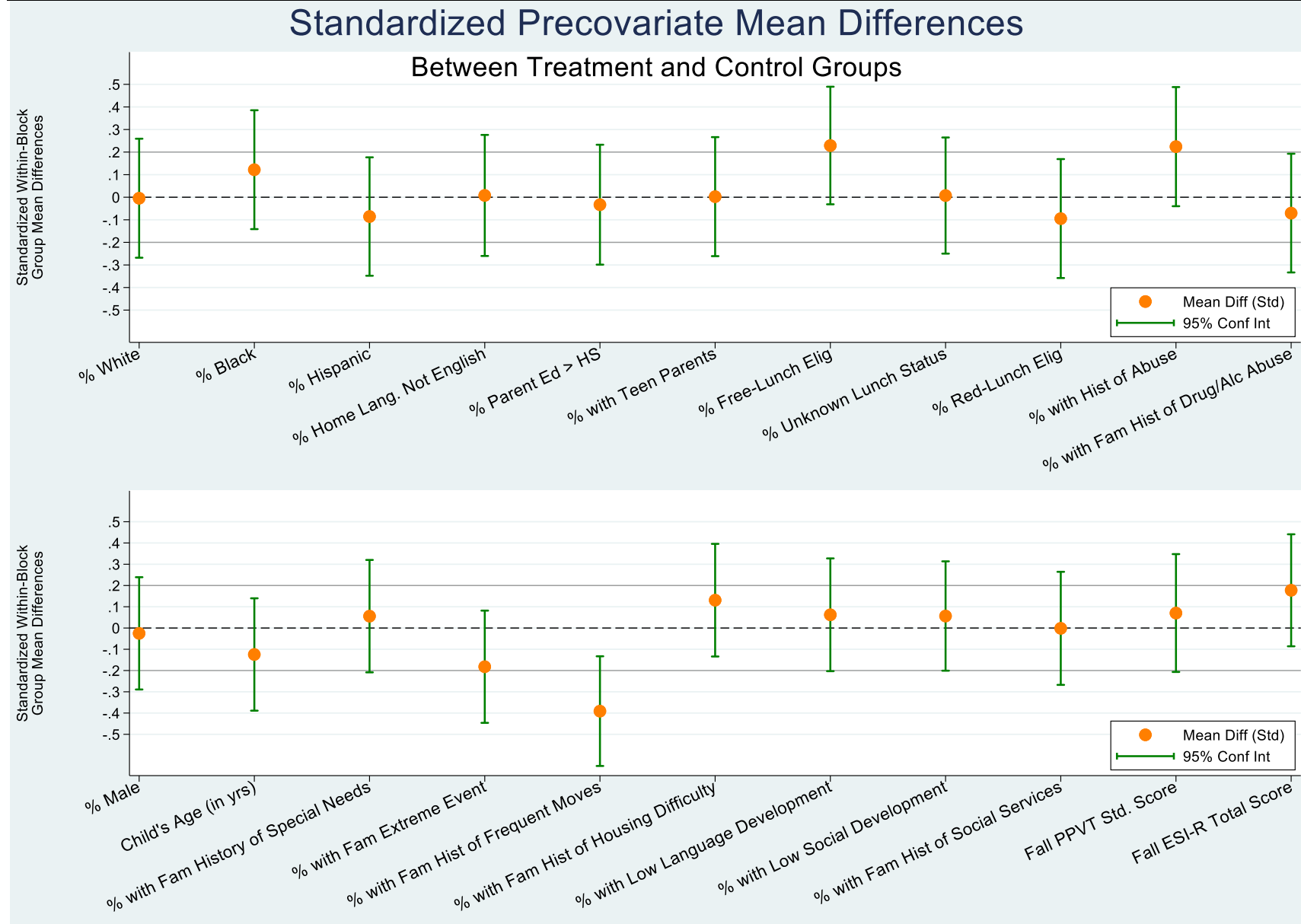
**Table 6. Primary and Exploratory Outcomes: Causal Effects of Full-Day Pre-K, when Missing Outcomes Imputed (ITT vs. CATE).**

Intent-to-Treat Analysis											
	Primary Outcomes		End of Pre-K TS GOLD						Fall Kindergarten DIBELS		
	<i>PPVT</i>	<i>ESI-R</i>	<i>Cognition</i>	<i>Language</i>	<i>Literacy</i>	<i>Math</i>	<i>Physical</i>	<i>Socio-</i>	<i>Overall</i>	<i>1st Sound</i>	<i>Ltr Naming</i>
	<i>(M3)</i>	<i>(M3)</i>						<i>Emotional</i>			
			<i>(M3)</i>	<i>(M3)</i>	<i>(M3)</i>	<i>(M3)</i>	<i>(M3)</i>	<i>(M3)</i>	<i>(M3)</i>	<i>(M3)</i>	<i>(M3)</i>
Assigned to Full	0.218 *	0.048	0.171 *	0.069	0.262 **	0.156 *	0.161 *	0.108	0.122	0.136	0.096
	(0.100)	(0.140)	(0.078)	(0.083)	(0.084)	(0.078)	(0.079)	(0.070)	(0.139)	(0.151)	(0.143)
Constant	0.104	0.176	0.082	0.055	0.071	0.045	-0.014	-0.002	0.112	0.067	0.133
	(0.080)	(0.111)	(0.062)	(0.065)	(0.067)	(0.062)	(0.063)	(0.055)	(0.110)	(0.120)	(0.113)
R^2	0.669	0.407	0.758	0.731	0.770	0.783	0.743	0.803	0.446	0.408	0.392
Adj. R^2	0.588	0.263	0.699	0.666	0.714	0.730	0.680	0.755	0.311	0.265	0.244
N	226	226	226	226	226	226	226	226	226	226	226
Complier Average Treatment Effects (Two-Stage Least Squares)											
	Primary Outcomes		End of Pre-K TS GOLD						Fall Kindergarten DIBELS		
	<i>PPVT</i>	<i>ESI-R</i>	<i>Cognition</i>	<i>Language</i>	<i>Literacy</i>	<i>Math</i>	<i>Physical</i>	<i>Socio-</i>	<i>Overall</i>	<i>1st Sound</i>	<i>Ltr Naming</i>
	<i>(M3)</i>	<i>(M3)</i>						<i>Emotional</i>			
			<i>(M3)</i>	<i>(M3)</i>	<i>(M3)</i>	<i>(M3)</i>	<i>(M3)</i>	<i>(M3)</i>	<i>(M3)</i>	<i>(M3)</i>	<i>(M3)</i>
Attended Full-Day	0.338 *	0.074	0.265 *	0.107	0.406 **	0.242 *	0.250 *	0.167	0.189	0.210	0.149
	(0.156)	(0.217)	(0.119)	(0.127)	(0.124)	(0.118)	(0.123)	(0.107)	(0.214)	(0.233)	(0.222)
Constant	0.021	0.158	0.016	0.029	-0.029	-0.015	-0.075	-0.043	0.066	0.015	0.096
	(0.111)	(0.154)	(0.084)	(0.090)	(0.088)	(0.083)	(0.087)	(0.076)	(0.152)	(0.166)	(0.157)
N	226	226	226	226	226	226	226	226	226	226	226

*Note: In this table, we impute possible assessment scores for study participants who are missing these outcomes. For study participants without scores in the treatment group, we assume they would have scored at the mean of the control group (within the same school site). For study participants without scores in the control group, we assume they would have scored at the mean of the treatment group. Results are reported for M3 only, which includes first-choice school (i.e., block) fixed effects, student-level control variables, and start of school year pre-scores on all assessments (and, where relevant, the language and format in which these assessments were taken). The vector of student-level pre-treatment covariates includes gender, race/ethnicity, free/reduced price lunch eligibility, and age at the start of pre-k. To control for any possible family context differences we also include parental education, whether or not child had a teen parent, whether the home language is English, as well as indicators for (pre-treatment reports of) history of family drug/alcohol abuse, family special needs, frequent moving, physical abuse, interactions with social services, difficulties obtaining secure housing, and perceived child underdeveloped language or social skills prior to pre-k. We include missingness dummies in cases where respondents have missing pre-treatment covariates.*

## Figures

Figure 1. Baseline Covariate Balance, Full Randomized Sample



## Appendix A: Baseline Covariate Balance and Attrition

**Table A1. Baseline Covariate Balance Comparison: Full Sample vs. Sample with End-of-Preschool TS-GOLD Exploratory Outcomes (Administered by District)**

Pre-Treatment Covariate	Treatment Group				Control Group			
	Full Sample		Non-Attrition		Full Sample		Non-Attrition	
	Mean	(N)	Mean	(N)	Mean	(N)	Mean	(N)
% White	11.4%	(114)	12.0%	(101)	11.6%	(112)	12.3%	(81)
% Black	0.8%	(114)	1.0%	(101)	0.0%	(112)	0.0%	(81)
% Hispanic	72.0%	(114)	72.7%	(101)	75.8%	(112)	79.7%	(81)
% Home Lang. Not English	48.8%	(112)	48.8%	(101)	48.4%	(102)	47.8%	(79)
% Parent Ed > HS	36.4%	(114)	37.5%	(101)	38.0%	(112)	47.1%	(81)
% with Teen Parents	1.8%	(114)	2.1%	(101)	1.8%	(112)	2.4%	(81)
% Free-Lunch Elig	66.2%	(114)	63.0%	(101)	55.0%	(112)	62.1%	(81)
% Unknown Lunch Status	21.8%	(114)	23.0%	(101)	21.5%	(112)	10.8%	(81)
% Red-Lunch Elig	11.3%	(114)	12.9%	(101)	14.4%	(112)	18.5%	(81)
% Male	48.0%	(114)	49.8%	(101)	49.3%	(112)	51.5%	(81)
Child's Age (in yrs)	4.34	(114)	4.36	(101)	4.38	(109)	4.38	(80)
% with Hist of Abuse	1.9%	(114)	2.1%	(101)	-0.2%	(112)	-0.2%	(81)
% with Fam Hist of Drug/Alc Abuse	0.9%	(114)	1.1%	(101)	1.7%	(112)	2.3%	(81)
% with Fam History of Special Needs	17.9%	(114)	19.0%	(101)	15.8%	(112)	19.5%	(81)
% with Fam Extreme Event	0.9%	(114)	1.1%	(101)	3.6%	(112)	4.8%	(81)
% with Fam Hist of Frequent Moves	4.6%	(114)	3.7%	(101)	16.7%	(112)	21.3%	(81)
% with Fam Hist of Housing Difficulty	4.7%	(114)	4.2%	(101)	2.3%	(112)	3.4%	(81)
% with Low Language Development	24.3%	(114)	25.5%	(101)	21.7%	(112)	25.0%	(81)
% with Low Social Development	38.1%	(114)	38.4%	(101)	35.4%	(112)	38.6%	(81)
% with Fam Hist of Social Services	2.6%	(114)	3.0%	(101)	2.7%	(112)	3.6%	(81)
Fall PPVT Std. Score	0.031	(105)	0.038	(98)	-0.040	(95)	-0.039	(73)
Fall ESI-R Total Score	0.084	(111)	0.140	(101)	-0.093	(104)	-0.111	(79)

*Note: "Full" sample refers all study participants originally assigned to each treatment group. The "Non-Attrition" sample refers to study participants who are observed with TS GOLD test scores in the spring of pre-k.*

**Table A2. Baseline Covariate Balance Comparison: Full Sample vs. Sample with Fall of Kindergarten DIBELS Exploratory Outcomes (Administered by District)**

Pre-Treatment Covariate	Treatment Group				Control Group			
	Full Sample		Non-Attrition		Full Sample		Non-Attrition	
	Mean	(N)	Mean	(N)	Mean	(N)	Mean	(N)
% White	11.4%	(114)	10.0%	(78)	11.6%	(112)	11.8%	(61)
% Black	0.8%	(114)	0.0%	(78)	0.0%	(112)	0.0%	(61)
% Hispanic	72.0%	(114)	75.7%	(78)	75.8%	(112)	78.6%	(61)
% Home Lang. Not English	48.8%	(112)	47.4%	(78)	48.4%	(102)	51.8%	(58)
% Parent Ed > HS	36.4%	(114)	36.5%	(78)	38.0%	(112)	38.6%	(61)
% with Teen Parents	1.8%	(114)	2.5%	(78)	1.8%	(112)	1.8%	(61)
% Free-Lunch Elig	66.2%	(114)	63.6%	(78)	55.0%	(112)	54.7%	(61)
% Unknown Lunch Status	21.8%	(114)	22.7%	(78)	21.5%	(112)	18.6%	(61)
% Red-Lunch Elig	11.3%	(114)	12.4%	(78)	14.4%	(112)	20.2%	(61)
% Male	48.0%	(114)	48.7%	(78)	49.3%	(112)	54.1%	(61)
Child's Age (in yrs)	4.34	(114)	4.34	(78)	4.38	(109)	4.33	(59)
% with Hist of Abuse	1.9%	(114)	1.3%	(78)	-0.2%	(112)	0.0%	(61)
% with Fam Hist of Drug/Alc Abuse	0.9%	(114)	0.1%	(78)	1.7%	(112)	1.5%	(61)
% with Fam History of Special Needs	17.9%	(114)	22.2%	(78)	15.8%	(112)	12.6%	(61)
% with Fam Extreme Event	0.9%	(114)	1.2%	(78)	3.6%	(112)	5.0%	(61)
% with Fam Hist of Frequent Moves	4.6%	(114)	6.3%	(78)	16.7%	(112)	19.8%	(61)
% with Fam Hist of Housing Difficulty	4.7%	(114)	5.2%	(78)	2.3%	(112)	-0.1%	(61)
% with Low Language Development	24.3%	(114)	22.3%	(78)	21.7%	(112)	24.0%	(61)
% with Low Social Development	38.1%	(114)	34.9%	(78)	35.4%	(112)	39.0%	(61)
% with Fam Hist of Social Services	2.6%	(114)	2.5%	(78)	2.7%	(112)	3.3%	(61)
Fall PPVT Std. Score	0.031	(105)	0.077	(75)	-0.040	(95)	-0.210	(55)
Fall ESI-R Total Score	0.084	(111)	0.211	(78)	-0.093	(104)	-0.002	(59)

*Note: "Full" sample refers all study participants originally assigned to each treatment group. The "Non-Attrition" sample refers to study participants who are observed with DIBELS test scores in the fall of pre-k.*

## References

- Barnett, W. S., Friedman-Krauss, A., Weisenfeld, G., Horowitz, M., Kasmin, R., & Squires, J. (2017). The state of preschool 2016: State preschool yearbook. New Brunswick, NJ: National Institute for Early Education Research. In: National Institute for Early Education Research New Brunswick, NJ.
- Bassok, D. (2010). Do Black and Hispanic children benefit more from preschool? Understanding differences in preschool effects across racial groups. *Child Development*, 81(6), 1828-1845.
- Bassok, D., Gibbs, C. R., & Latham, S. (2018). Preschool and Children's Outcomes in Elementary School: Have Patterns Changed Nationwide Between 1998 and 2010? *Child Development*.
- Bassok, D., & Loeb, S. (2015). Early childhood and the achievement gap. *Handbook of Research in Education Finance and Policy*, 510-528.
- Belsky, J. (2002). Quantity counts: Amount of child care and children's socioemotional development. *Journal of Developmental & Behavioral Pediatrics*, 23(3), 167-170.
- Blau, D., & Currie, J. (2006). Pre-school, day care, and after-school care: who's minding the kids? *Handbook of the Economics of Education*, 2, 1163-1278.
- Camilli, G., Vargas, S., Ryan, S., & Barnett, W. S. (2010). Meta-analysis of the effects of early education interventions on cognitive and social development. *Teachers College Record*, 112(3), 579-620.
- Campbell, F. A., & Ramey, C. T. (1994). Effects of early intervention on intellectual and academic achievement: a follow-up study of children from low-income families. *Child Development*, 65(2), 684-698.
- Campbell, F. A., & Ramey, C. T. (1995). Cognitive and school outcomes for high-risk African-American students at middle adolescence: Positive effects of early intervention. *American Educational Research Journal*, 32(4), 743-772.
- Cannon, J. S., Jacknowitz, A., & Painter, G. (2006). Is full better than half? Examining the longitudinal effects of full-day kindergarten attendance. *Journal of Policy Analysis and Management: The Journal of the Association for Public Policy Analysis and Management*, 25(2), 299-321.
- Child\_Trends\_Data\_Bank. (2015). *Full-day kindergarten: Indicators on children and youth*. Retrieved from Child Trends Data Bank. : Available at: <http://www.childtrends.org>
- DeCicca, P. (2007). Does full-day kindergarten matter? Evidence from the first two years of schooling. *Economics of Education Review*, 26(1), 67-82.
- Deming, D. (2009). Early childhood intervention and life-cycle skill development: Evidence from Head Start. *American Economic Journal: Applied Economics*, 1(3), 111-134.
- Gibbs, C. (2014). Experimental evidence on early intervention: The impact of full-day kindergarten. *Batten School of Leadership and Public Policy, University of Virginia* [http://www.batten.virginia.edu/sites/default/files/fwpapers/Gibbs\\_full-day% 20K% 20experiment. pdf](http://www.batten.virginia.edu/sites/default/files/fwpapers/Gibbs_full-day%20K%20experiment.pdf).
- Gormley, W. T. (2008). The effects of Oklahoma's pre-k program on Hispanic children. *Social Science Quarterly*, 89(4), 916-936.
- Gullo, D. F. (2000). The long term educational effects of half-day vs full-day kindergarten. *Early Child Development and Care*, 160(1), 17-24.
- Heckman, J. J. (2006). Skill formation and the economics of investing in disadvantaged children. *Science*, 312(5782), 1900-1902.

- Kena, G., Hussar, W., McFarland, J., de Brey, C., Musu-Gillette, L., Wang, X., . . . Diliberti, M. (2016). The Condition of Education 2016. NCES 2016-144. *National Center for Education Statistics*.
- Leak, J., Duncan, G. J., Li, W., Magnuson, K., Schindler, H., & Yoshikawa, H. (2010). *Is timing everything? How early childhood education program impacts vary by starting age, program duration and time since the end of the program*. Paper presented at the Biennial Meeting for the Society for Research on Child Development, Montreal, Quebec, Canada.
- Lee, V. E., Burkam, D. T., Ready, D. D., Honigman, J., & Meisels, S. J. (2006). Full-day versus half-day kindergarten: In which program do children learn more? *American Journal of Education*, 112(2), 163-208.
- 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*.
- Lipsey, M. W., Hofer, K. G., Dong, N., Farran, D. C., & Bilbrey, C. (2013). Evaluation of the Tennessee Voluntary Prekindergarten Program: Kindergarten and First Grade Follow-Up Results from the Randomized Control Design. Research Report. *Peabody Research Institute*.
- Loeb, S., Bridges, M., Bassok, D., Fuller, B., & Rumberger, R. W. (2007). How much is too much? The influence of preschool centers on children's social and cognitive development. *Economics of Education Review*, 26(1), 52-66.
- Magnuson, K. A., Ruhm, C., & Waldfogel, J. (2007). The persistence of preschool effects: Do subsequent classroom experiences matter? *Early Childhood Research Quarterly*, 22(1), 18-38.
- Philips, D. A., Lipsey, M. W., Dodge, K. A., Haskins, R., Bassok, D., Burchinal, M., . . . Weiland, C. (2017). *Puzzling It Out: The Current State of Scientific Knowledge on Pre-Kindergarten Effects: A Consensus Statement*. Retrieved from [https://www.brookings.edu/wp-content/uploads/2017/04/consensus-statement\\_final.pdf](https://www.brookings.edu/wp-content/uploads/2017/04/consensus-statement_final.pdf)
- Puma, M., Bell, S., Cook, R., Heid, C., Broene, P., Jenkins, F., & Downer, J. (2012). Third grade follow-up to the Head Start Impact Study: Final report (OPRE Report 2012-45). Washington, DC: US Department of Health and Human Services, Administration for Children and Families, Office of Planning. *Research and Evaluation*, 28.
- Shonkoff, J. P., & Phillips, D. A. (2000). *From neurons to neighborhoods: The science of early childhood development*: National Academies Press.
- Vandell, D. L., Belsky, J., Burchinal, M., Steinberg, L., & Vandergrift, N. (2010). Do effects of early child care extend to age 15 years? Results from the NICHD study of early child care and youth development. *Child Development*, 81(3), 737-756.
- Votruba-Drzal, E., Li-Grining, C. P., & Maldonado-Carreño, C. (2008). A Developmental Perspective on Full-Versus Part-Day Kindergarten and Children's Academic Trajectories Through Fifth Grade. *Child Development*, 79(4), 957-978.
- Walters, C. R. (2015). Inputs in the production of early childhood human capital: Evidence from Head Start. *American Economic Journal: Applied Economics*, 7(4), 76-102.
- Weiland, C. (2018). Pivoting to the "how": Moving preschool policy, practice, and research forward. *Early Childhood Research Quarterly*.
- Weiland, C., & Yoshikawa, H. (2013). Impacts of a prekindergarten program on children's mathematics, language, literacy, executive function, and emotional skills. *Child Development*, 84(6), 2112-2130.

- Whitehurst, G. J. (2016). Family Support or School Readiness? Contrasting Models of Public Spending on Children's Early Care and Learning. Evidence Speaks Reports, Vol 1, # 16. *Center on Children and Families at Brookings*.
- Wong, V. C., Cook, T. D., Barnett, W. S., & Jung, K. (2008). An effectiveness-based evaluation of five state pre-kindergarten programs. *Journal of Policy Analysis and Management*, 27(1), 122-154.
- Zvoch, K., Reynolds, R. E., & Parker, R. P. (2008). Full-day kindergarten and student literacy growth: Does a lengthened school day make a difference? *Early Childhood Research Quarterly*, 23(1), 94-107.
- Abry, T., Latham, S., Bassok, D., & LoCasale-Crouch, J. (2015). Preschool and kindergarten teachers' beliefs about early school competencies: Misalignment matters for kindergarten adjustment. *Early Childhood Research Quarterly*, 31, 78-88.
- Araujo, M. C., Carneiro, P., Cruz-Aguayo, Y., & Schady, N. (2016). Teacher quality and learning outcomes in kindergarten. *The Quarterly Journal of Economics*, 131(3), 1415-1453.
- Barnett, W. S., Friedman-Krauss, A., Weisenfeld, G., Horowitz, M., Kasmin, R., & Squires, J. (2017). The state of preschool 2016: State preschool yearbook. New Brunswick, NJ: National Institute for Early Education Research. In: National Institute for Early Education Research New Brunswick, NJ.
- Barnett, W. S., Jung, K., Friedman-Krauss, A., Frede, E. C., Nores, M., Hustedt, J. T., . . . Daniel-Echols, M. (2018). State Prekindergarten Effects on Early Learning at Kindergarten Entry: An Analysis of Eight State Programs. *AERA Open*, 4(2), 2332858418766291.
- Bassok, D. (2010). Do Black and Hispanic children benefit more from preschool? Understanding differences in preschool effects across racial groups. *Child Development*, 81(6), 1828-1845.
- Bassok, D., Fitzpatrick, M., Greenberg, E., & Loeb, S. (2016). Within-and between-sector quality differences in early childhood education and care. *Child Development*, 87(5), 1627-1645.
- Bassok, D., Gibbs, C. R., & Latham, S. (2018). Preschool and Children's Outcomes in Elementary School: Have Patterns Changed Nationwide Between 1998 and 2010? *Child Development*.
- Belsky, J. (2002). Quantity counts: Amount of child care and children's socioemotional development. *Journal of Developmental & Behavioral Pediatrics*, 23(3), 167-170.
- Blau, D., & Currie, J. (2006). Pre-school, day care, and after-school care: who's minding the kids? *Handbook of the Economics of Education*, 2, 1163-1278.
- Brownell, M. D., Nickel, N. C., Chateau, D., Martens, P. J., Taylor, C., Crockett, L., . . . Goh, C. Y. (2015). Long-term benefits of full-day kindergarten: a longitudinal population-based study. *Early Child Development and Care*, 185(2), 291-316.
- Burke, M. D., Hagan-Burke, S., Kwok, O., & Parker, R. (2009). Predictive validity of early literacy indicators from the middle of kindergarten to second grade. *The Journal of Special Education*, 42(4), 209-226.
- Campbell, F. A., Pungello, E. P., Burchinal, M., Kainz, K., Pan, Y., Wasik, B. H., . . . Ramey, C. T. (2012). Adult outcomes as a function of an early childhood educational program: an Abecedarian Project follow-up. *Developmental psychology*, 48(4), 1033.
- Campbell, F. A., & Ramey, C. T. (1994). Effects of early intervention on intellectual and academic achievement: a follow-up study of children from low-income families. *Child Development*, 65(2), 684-698.

- Cannon, J. S., Jacknowitz, A., & Painter, G. (2006). Is full better than half? Examining the longitudinal effects of full-day kindergarten attendance. *Journal of Policy Analysis and Management: The Journal of the Association for Public Policy Analysis and Management*, 25(2), 299-321.
- Child\_Trends\_Data\_Bank. (2015). *Full-day kindergarten: Indicators on children and youth*. Retrieved from Child Trends Data Bank. : Available at: <http://www.childtrends.org>
- Clements, D. H., & Sarama, J. (2008). Experimental evaluation of the effects of a research-based preschool mathematics curriculum. *American Educational Research Journal*, 45(2), 443-494.
- Cooper, H., Allen, A. B., Patall, E. A., & Dent, A. L. (2010). Effects of full-day kindergarten on academic achievement and social development. *Review of Educational Research*, 80(1), 34-70.
- DeCicca, P. (2007). Does full-day kindergarten matter? Evidence from the first two years of schooling. *Economics of Education Review*, 26(1), 67-82.
- Dunn, L., & Dunn, D. (2013). PPVT-4 Technical Report. In: Minneapolis, MN: Pearson Assessments.
- Dunn, L. M., & Dunn, D. M. (2007). *PPVT-4: Peabody picture vocabulary test. Form A*: Pearson Assessments.
- Early, D. M., Maxwell, K. L., Ponder, B. D., & Pan, Y. (2017). Improving teacher-child interactions: A randomized controlled trial of Making the Most of Classroom Interactions and My Teaching Partner professional development models. *Early Childhood Research Quarterly*, 38, 57-70.
- Gibbs, C. (2014). Experimental evidence on early intervention: The impact of full-day kindergarten. *Batten School of Leadership and Public Policy, University of Virginia* [http://www.batten.virginia.edu/sites/default/files/fwpapers/Gibbs\\_full-day% 20K% 20experiment.pdf](http://www.batten.virginia.edu/sites/default/files/fwpapers/Gibbs_full-day%20K%20experiment.pdf).
- Gibbs, C. R. (2014). Experimental evidence on early intervention: The impact of full-day kindergarten. *Frank Batten School of Leadership and Public Policy Working Paper*, 4.
- Good, R., & Kaminski, R. (2002). *Dynamic Indicators of Basic Early Literacy Skills (DIBELS)* (Vol. 6). Eugene, OR: Institute for the Development of Education Achievement. Available: <http://dibels.uoregon.edu>.
- Good, R., Kaminski, R., Shinn, M., Bratten, J., Shinn, M., Laimon, D., & Flindt, N. (2004). Technical adequacy of DIBELS: Results of the Early Childhood Research Institute on measuring growth and development. *Eugene, OR: University of Oregon*.
- Gormley, W. T. (2008). The effects of Oklahoma's pre-k program on Hispanic children. *Social Science Quarterly*, 89(4), 916-936.
- Gullo, D. F. (2000). The long term educational effects of half-day vs full-day kindergarten. *Early Child Development and Care*, 160(1), 17-24.
- Head Start Performance Standards*. (2016). Washington, DC: Federal Register Vol 81, Issue 172, Sept 6, 2016 Retrieved from <https://www.gpo.gov/fdsys/pkg/FR-2016-09-06/pdf/2016-19748.pdf>.
- Heckman, J. J. (2006). Skill formation and the economics of investing in disadvantaged children. *Science*, 312(5782), 1900-1902.
- Heroman, C., Burts, D. C., Berke, K.-I., Bickart, T. S., Nelson, H. P., Taub, L., & Boyle, K. (2010). Objectives for Development & Learning: Birth Through Kindergarten.



- Herry, Y., Maltais, C., & Thompson, K. (2007). Effects of a full-day preschool program on 4-year-old children. *Early Childhood Research & Practice*, 9(2), n2.
- Kena, G., Hussar, W., McFarland, J., de Brey, C., Musu-Gillette, L., Wang, X., . . . Diliberti, M. (2016). The Condition of Education 2016. NCES 2016-144. *National Center for Education Statistics*.
- Kim, D.-H., Lambert, R. G., & Burts, D. C. (2013). Evidence of the validity of Teaching Strategies GOLD® assessment tool for English language learners and children with disabilities. *Early Education & Development*, 24(4), 574-595.
- Lee, V. E., Burkam, D. T., Ready, D. D., Honigman, J., & Meisels, S. J. (2006). Full-day versus half-day kindergarten: In which program do children learn more? *American Journal of Education*, 112(2), 163-208.
- Leow, C., & Wen, X. (2017). Is full day better than half day? A propensity score analysis of the association between Head Start Program intensity and children's school performance in kindergarten. *Early Education and Development*, 28(2), 224-239.
- Loeb, S., Bridges, M., Bassok, D., Fuller, B., & Rumberger, R. W. (2007). How much is too much? The influence of preschool centers on children's social and cognitive development. *Economics of Education Review*, 26(1), 52-66.
- Meisels, S. J., Henderson, L. W., Liaw, F.-r., Browning, K., & Ten Have, T. (1993). New evidence for the effectiveness of the Early Screening Inventory. *Early Childhood Research Quarterly*, 8(3), 327-346.
- Meisels, S. J., Marsden, D. B., Wiske, M. S., & Henderson, L. W. (1997). *ESI-R: Early Screening Inventory-Revised. Examiner's Manual*: ERIC.
- Moodie, S., Daneri, P., Goldhagen, S., Halle, T., Green, K., & LaMonte, L. (2014). Early Childhood Developmental Screening: A Compendium of Measures for Children Ages Birth to Five. OPRE Report 2014-11. *US Department of Health and Human Services*.
- Morris, P. A., Connors, M., Friedman-Krauss, A., McCoy, D. C., Weiland, C., Feller, A., . . . Yoshikawa, H. (2018). New findings on impact variation from the Head Start Impact Study: Informing the scale-up of early childhood programs. *AERA Open*, 4(2), 2332858418769287.
- Philips, D. A., Lipsey, M. W., Dodge, K. A., Haskins, R., Bassok, D., Burchinal, M., . . . Weiland, C. (2017). *Puzzling It Out: The Current State of Scientific Knowledge on Pre-Kindergarten Effects: A Consensus Statement*. Retrieved from [https://www.brookings.edu/wp-content/uploads/2017/04/consensus-statement\\_final.pdf](https://www.brookings.edu/wp-content/uploads/2017/04/consensus-statement_final.pdf)
- Pianta, R. C., La Paro, K. M., & Hamre, B. K. (2008). *Classroom Assessment Scoring System™: Manual K-3*: Paul H Brookes Publishing.
- Piasta, S. B., Justice, L. M., O'Connell, A. A., Mauck, S. A., Weber-Mayrer, M., Schachter, R. E., . . . Spear, C. F. (2017). Effectiveness of large-scale, state-sponsored language and literacy professional development on early childhood educator outcomes. *Journal of Research on Educational Effectiveness*, 10(2), 354-378.
- Puma, M., Bell, S., Cook, R., Heid, C., & Lopez, M. (2005). Head Start Impact Study: First Year Findings. *Administration for Children & Families*.
- Reynolds, A. J., Richardson, B. A., Hayakawa, M., Lease, E. M., Warner-Richter, M., Englund, M. M., . . . Sullivan, M. (2014). Association of a full-day vs part-day preschool intervention with school readiness, attendance, and parent involvement. *Jama*, 312(20), 2126-2134.
- Robin, K. B., Frede, E. C., & Barnett, W. S. (2006). Is more better? The effects of full-day vs half-day preschool on early school achievement.

- Rouse, H. L., & Fantuzzo, J. (2006). Validity of the Dynamic Indicators for Basic Early Literacy Skills as an indicator of early literacy for urban kindergarten children. *School Psychology Review, 35*(3), 341.
- 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.
- Shonkoff, J. P., & Phillips, D. A. (2000). *From neurons to neighborhoods: The science of early childhood development*. National Academies Press.
- Snow, C. E., & Matthews, T. J. (2016). Reading and language in the early grades. *The Future of Children, 57*-74.
- Teaching Strategies. (2011). *Teaching Strategies GOLD Assessment System Technical Summary*. Retrieved from
- Teaching Strategies. (2013). *Teaching strategies GOLD assessment system: Concurrent validity*. Retrieved from
- Vandell, D. L., Belsky, J., Burchinal, M., Steinberg, L., & Vandergrift, N. (2010). Do effects of early child care extend to age 15 years? Results from the NICHD study of early child care and youth development. *Child Development, 81*(3), 737-756.
- Votruba-Drzal, E., Li-Grining, C. P., & Maldonado-Carreño, C. (2008). A Developmental Perspective on Full-Versus Part-Day Kindergarten and Children's Academic Trajectories Through Fifth Grade. *Child Development, 79*(4), 957-978.
- Walters, C. R. (2015). Inputs in the production of early childhood human capital: Evidence from Head Start. *American Economic Journal: Applied Economics, 7*(4), 76-102.
- Weiland, C. (2018). Pivoting to the "how": Moving preschool policy, practice, and research forward. *Early Childhood Research Quarterly*.
- Weiland, C., & Yoshikawa, H. (2013). Impacts of a prekindergarten program on children's mathematics, language, literacy, executive function, and emotional skills. *Child Development, 84*(6), 2112-2130.
- 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.
- Whitebook, M., Phillips, D., & Howes, C. (2014). Worthy work, STILL unlivable wages: The early childhood workforce 25 years after the National Child Care Staffing Study. *Center for the Study of Child Care Employment, University of California, Berkeley*.
- Wong, V. C., Cook, T. D., Barnett, W. S., & Jung, K. (2008). An effectiveness-based evaluation of five state pre-kindergarten programs. *Journal of Policy Analysis and Management, 27*(1), 122-154.
- Zvoch, K., Reynolds, R. E., & Parker, R. P. (2008). Full-day kindergarten and student literacy growth: Does a lengthened school day make a difference? *Early Childhood Research Quarterly, 23*(1), 94-107.

---

<sup>1</sup> In July 2016 alone, more than twice as many families applied for FDPK than could be accommodated.

<sup>2</sup> The exact wording of the language development question was: “Is your child in need of language development including, but not limited to, the ability to speak English?”. The exact wording of the question about social development was: “Does your child have problems with social situations?”. See footnote of Table 2 for the wording of all reported questions for parents on the general application.