HHS Public Access

Author manuscript

Child Youth Serv Rev. Author manuscript; available in PMC 2020 June 01.

Published in final edited form as:

Child Youth Serv Rev. 2019 June ; 101: 285–298. doi:10.1016/j.childyouth.2019.04.005.

Dosage Effects in the Child-Parent Center PreK-to-3rd Grade Program: A Re-Analysis in the Chicago Longitudinal Study

Suh-Ruu Ou, PhD[1], **Irma Arteaga, PhD**[2], and **Arthur Reynolds, PhD**[1]

1. Institute of Child Development, University of Minnesota, Minneapolis, MN

2. Harry Truman School of Public Affairs, University of Missouri, Columbia, MO

Abstract

Although substantial investments in early childhood intervention have continued, whether gains are sustained past kindergarten for routinely implemented programs is a critical research need. Using data from the Chicago Longitudinal Study (CLS; N=1,539; 50.3% female; 92.9% African American and 7.1% Hispanic), an on-going investigation of the Child-Parent Center (CPC) program for an inner-city cohort, this study investigates the effects of program duration from preschool to 3rd grade on school outcomes and whether the effects differ by gender. Regression analyses are conducted to compare the differences in outcomes among intervention groups. Inverse probability weighting (IPW) is used to adjust for potential attrition and selection biases. Findings indicate that relative to the preschool plus kindergarten (P-K) group, participation from preschool through third grade (P–3) is significantly associated with better academic functioning at both 3rd and 8th grades, better classroom adjustment at 3rd grade, lower rates of retention and school mobility, and few years of special education. Relative to the preschool through second grade (P–2) group, the P–3 group has significantly higher academic functioning in third grade. Results suggest that the P–3 dosage is associated with larger effects on academic functioning for girls and larger effects on social-emotional functioning for boys compare to the P-K dosage. Findings suggest that receiving up to third grade (P-3) of an early childhood education program have associated with persistent effects on developmental outcomes compared to the dosages of P-K. Multi-year programs have the potential to sustain early childhood gains and promote healthy development via improving academic functioning and school experiences.

Keywords

Early childhood intervention; Early Childhood Education Program; Preschool through Third-Grade intervention; Child-Parent Center; dosage effects; gender differences

Corresponding Author: Suh-Ruu Ou, Institute of Child Development, University of Minnesota, 51 E. River Road, Minneapolis, MN 55455, (sou@umn.edu), 612-626-0446.

Declarations of interest: None.

1. Introduction

Early Childhood Education (ECE) intervention is recognized as an effective approach to narrow the achievement gaps by income and race/ethnicity (Burger, 2010). ECE can also serve as a policy lever to improve public health by improving social-emotional functioning in early childhood (Moore et al., 2015) and by reducing the growing health disparities resulting from variability in adulthood educational attainment (Muennig, 2015). Thus, ECE intervention is essential to improve health, human capital, and well-being across the life course (Black et al., 2017; Conti, Heckman, & Pinto, 2015). How gains from ECE are initiated and sustained have received increased attention because recent findings on the short-term effects of public ECE programs dissipate when children enter elementary school.

1.1 Background

Early childhood landscape has changed considerably since 1965, and there are several points worth noting. First, the percentage of 3–5 year old children enrolled in ECE programs have increased significantly (See Appendix A). The increase was rapid between 1965 and 1983 for all children ages 3, 4, and 5; the increase continued but slowed down between 1983 and 1997. The percentages of children enrolled in ECE programs have been stable since 1997. Children from economically disadvantaged families receive priority for enrollment in many public pre-kindergarten programs, but the disparities in access to preschool by family income, race and ethnicity remain (Chaudry & Datta, 2017). Second, resources on ECE have increased significantly and there is more variation in funding sources. Federal and state investments in preschool have increased in the last decade. For example, the Preschool Development Grants awarded $463 million and state matching funds of $985 million in 2014 and 2015 (U.S. Department of Education, 2015) to provide children access to preschool. States spending on preschool rose from just $2.4 billion in 2002 to over $7.6 billion in 2017 (National Institute of Early Education Research (NIEER), 2018). Many preschools combine private with public dollars, and different sources of public dollars, such as Head Start, Child-care subsidies, and public pre-kindergarten funds, are combined as well (Chaudry & Datta, 2017). Finally, although the early childhood landscape has changed since 1965, the positive effects of preschool in academic outcomes have remained consistent over the past decades (Camilli, Vargas, Ryan, & Barnett, 2010; Cannon et al., 2017a; Karoly, Kilburn, & Cannon, 2005). For example, a recent study reported that preschool is positively associated with academic outcomes using two nationally representative U.S. cohorts 12 years apart from each other (Early Childhood Longitudinal Study, ECLS–K 1998 and ECLS–K 2010) (Bassok, Gibbs, & Latham, 2018). While the landscape of the ECE has changed, findings from programs implemented decades ago can still provide insights on how gains from ECE are initiated and sustained.

1.1.1 Effects of ECE programs.—

Effects of ECE programs were examined in numerous studies. Short-term effects of preschool programs on school readiness are reported consistently in reviews and meta-analyses (Blok, Fukkink, Gebhardt, & Leseman, 2005; Burger, 2010; Camilli et al., 2010; Farran, 2000; Yoshikawa et al., 2013). However, there is limited evidence on the long-term effects of public programs (Karoly et al., 2005; Phillips et al., 2017). A meta-analysis of state prekindergarten programs by Gilliam and Zigler (2000)

found that the most significant impacts were limited to kindergarten and first grade. More recent evidence, however, shows that gains can be sustained into middle school and beyond (Cannon et al., 2017b; McCoy et al., 2017; Meloy, Gardner, & Darling-Hammond, 2019). Findings of mixed (Cascio & Schanzenbach, 2013; Hill, Gormley, & Adelstein, 2015; Jung, Barnett, Hustedt, & Francis, 2013) and unsustained effects (Lipsey, Farran, & Hofer, 2015; Puma et al., 2012) raise concerns about the effectiveness of scale up efforts. Only a few studies of contemporary state prekindergarten programs have reported enduring effects into middle school (Barnett, Jung, Youn, & Frede, 2013; Dodge, Bai, Ladd, & Muschkin, 2017; Gormley, Phillips, & Anderson, 2018) and beyond (Schweinhart, Xiang, Daniel-Echols, Browning, & Wakabayashi, 2012). How gains from ECE programs are initiated, increased, and sustained continues to be a critical issue in the field.

Bailey and colleagues (2017) discussed three processes, skill-building, foot-in-the-door, and sustaining environments, that might explain whether the impacts of ECE programs persist or fade out. ECE quality itself was not addressed as an additional explanation. Skill-building and foot-in-the-door emphasize the right kinds of skills and capacities equip children to take advantage of an environmental opportunity or skill development while sustaining environments emphasizes the necessity of subsequent investments in sufficiently high-quality schools and other environmental contexts for persistent effects (Bailey et al., 2017). Sustaining environments is especially relevant because the fade out of positive effects has been attributed to the low quality of elementary school children attended subsequently (Benner, Thornton, & Crosnoe, 2016; Currie & Thomas, 2000; Phillips et al., 2017). Qualities of elementary school that are related to the persistence of preschool effects include the higher quality classroom environments during the elementary school years (Ansari & Pianta, 2018a, 2018b), attendance of high performing school (Zhai, Raver, & Jones, 2012), consistency of instructional practices (Mashburn & Yelverton, 2019), better teaching quality in early grades (Swain, Springer, & Hofer, 2015), and small class size and high levels of reading instruction in kindergarten (Magnuson, Ruhm, & Waldfogel, 2007).

For example, the benefits of Head Start spending were found to be larger when followed by access to better-funded public K–12 schools, and the increases in K–12 spending were more efficacious for poor children who were exposed to higher levels of Head Start spending during their preschool years (Johnson & Jackson, 2017). Elementary school quality indicators that are not related to persistence of preschool effects includes academic content coverage in kindergarten (Claessens, Engel, & Curran, 2014). Targeted teacher professional supports were found to mitigate fadeout between kindergarten and first grade but it was not mediated through classroom quality (Jenkins et al., 2018). It should be noted, however, that the contributions of all of the above processes and resource supports are likely to vary as a function of the quality of the preschool experience and the size of the initial effect (Reynolds, Ou, Mondi, & Hayakawa, 2017; Reynolds & Temple, 2019).

Given the overall evidence of both sustained and unsustained gains of large-scale ECE programs, it is crucial to learn from large-scale ECE programs that report positive outcomes in adulthood. Evidence on long-term effects of large-scale programs is sparse. Knowledge on the persistent impacts of ECE programs is timely given the focus of national and state efforts to increase access to ECE and pressing needs to understand how to sustain gains from

preschool to lead to long-term effects. The present study re-examines the effects of such a large-scale public program, Child-Parent Center (CPC) Program, from a new perspective in the hope to provide insights into the persistent and fadeout effects of large-scale ECE programs. As one of the few public ECE programs that have reported long-term effects, findings from the present study will have implications on sustained ECE effects after third grade and beyond.

1.2 Heterogeneous Effects by Gender

Although heterogeneous effects of ECE by race/ethnicity, dual language status, and low-income status have been evaluated (Duncan & Sojourner, 2013; Ladd, 2017; Yoshikawa et al., 2013), the heterogeneous effects of ECE by gender are under-investigated. Gender differences in education and behavior are well known. Studies have found that girls have more advanced reading skills, have advantages in social skills and classroom behavior, and obtain higher grades in school than boys (Buchmann, DiPrete, & McDaniel, 2008; Fortin, Oreopoulos, & Phipps, 2015; Loveless, 2015).

Schore (2017) uses a model of the psychoneurobiological mechanisms to underline the vulnerability of the developing male. Developing males are more vulnerable over a longer period of time to stressors in the social environment and toxins in the physical environment that negatively impact right-brain development because male brain matures more slowly than female brain in the prenatal, perinatal, and postnatal critical periods. Boys are more vulnerable early in life than girls, so they are more likely to be impacted by adverse life conditions than girls (Autor, Figlio, Karbownik, Roth, & Wasserman, 2015; Golding & Fitzgerald, 2017). This body of literature suggests that we might find substantial differences in the impacts of ECE by gender. In addition, the long-term effects of ECE might be promoted differently by gender. Examining boys and girls together ignores potentially large differences in treatment effects (García, Heckman, & Ziff, 2018). There is no consensus on whether ECE program impacts differ by gender in certain ways (García et al., 2018; Hill et al., 2015; Magnuson et al., 2016; Ou & Reynolds, 2006; Schweinhart et al., 2005). Nevertheless, given that ECE programs were found to benefit economically disadvantaged children more than their more advantaged peers (Ladd, 2017; Magnuson, Meyers, Ruhm, & Waldfogel, 2004), it is likely that boys might benefit more than girls because boys are more vulnerable early in life than girls. Gender differential effects of ECE are important from a developmental science perspective, as they will shed light on the mechanisms of ECE effects by gender and help to develop strategies to reduce gender gaps in education and health (Muschkin, Ladd, Dodge, & Bai, 2018).

1.3 The Present Study

The present study investigates whether the effects of the Child-Parent Center (CPC) program on academic functioning, social-emotional functioning in school, and school experiences between ages 9 and 16 differ by duration of participation and whether the effects differ by gender. Findings from the present study help identify a threshold of ECE dosage associated with positive effects in early adolescence and potentially leading to long-term effects.

1.3.1 The CPC program.—The CPC program had a unique history in that it not only provided comprehensive services similar to Head Start but was the first early education program funded under Title I of the Elementary and Secondary Education Act. Opened in 1967 in four disenfranchised neighborhoods in need of revitalization, the program showed strong effects early on and was expanded over the next decade. The CPC programs have been implemented in Chicago although it might not have received as much attention as other state prekindergarten programs. The program history and elements are described in other reports (Reynolds, 2000; Reynolds, Hayakawa, Candee, & Englund, 2016). A summary of the main features is provided here.

CPC is a federally-funded, enriched early childhood education intervention that serves children from preschool through third grade (ages 3–9). CPC targets Title I children and offers comprehensive services. The CPC preschool program is followed by a full-day or part-day CPC kindergarten. The CPC school-age program is open to any child in the school in either first through second grades in 14 sites or first through third grades in 6 sites. All participants attended either CPC kindergarten or an alternative kindergarten program. The CPC intervention underscores the acquisition of basic skills in language arts and math through both teacher- directed and child-initiated activities. All teachers have bachelor's degrees and are certified in early childhood education. Major components of the intervention (preschool, kindergarten, and school-age) include center-based education; instructional supports; small class sizes; a parent program that includes parenting education, parent room activities, classroom volunteering, and home visitation; and health and nutrition services, including screening and diagnostic services, meal services and referral by program nurses. Parents are expected to participate in the program up to half of a day per week through various supported activities. The program's focus on the continuity of learning environments indicates that optimal development can be promoted through enriched experiences and settings created together by children, families, and schools.

Previous studies have revealed significant benefits of the CPC preschool participation on multidimensional well-being (Reynolds, Temple, Ou, Arteaga, & White, 2011; Reynolds et al., 2007; Reynolds, Temple, Robertson, & Mann, 2001), the high economic returns of the CPC program (Reynolds, Temple, White, Ou, & Robertson, 2011), and the mechanisms and processes of change from cognitive and motivational advantages to enhancements in social-emotional development (Reynolds & Ou, 2011; Reynolds, Ou, & Topitzes, 2004). The CPC program is a Preschool through third-grade (P–3) program that provides children both preschool and school- age services up to third grade. To receive the full dosage of services, children participate in the P-3 program from preschool to third grade. This enhances continuity in learning, defined as the consistency and predictability in environmental settings (Reynolds, Hayakawa, et al., 2017). However, the effects of the CPC program have not been examined via the P-3 framework.

Effects of the CPC program have been examined separately by preschool and school-age program, and dosage of the CPC program has been examined by extended participation (4–6 years) and less extended participation (0–3 years) (Reynolds, Temple, Ou, et al., 2011; Reynolds et al., 2007; Reynolds et al., 2001). Dosage of CPC preschool have been examined by years of preschool participation (Arteaga, Humpage, Reynolds, & Temple, 2014;

Reynolds, 1995). Researchers suggest that providing children with continuity of service from preschool through third grade (P–3) will sustain the gains in preschool and lead to better developmental outcomes (Benner et al., 2016; Reynolds, Magnuson, & Ou, 2010). Yet empirical evidence on the value- added of preschool to 3rd grade (P–3) continuity is scant. The present study examines the dosage of the CPC program on the P-3 framework for the first time.

1.3.2 Theoretical perspectives of the CPC P-3 program.—The conceptual frameworks behind the CPC P–3 program are ecological systems theory (Bronfenbrenner, 1989) and psychosocial risk and protective factors (Rutter & Rutter, 1993). They specify that developmental outcomes are enhanced as a function of supportive and protective contexts and behaviors at multiple levels of influence. The concept of developmental continuity or environmental maintenance of development is a prominent feature, as the effects of early childhood experiences are magnified as they increase in duration and intensity (Ramey & Ramey, 2019). An additional component of the theories is a cumulative advantage and the similar concept of developmental cascades (Masten & Cicchetti, 2010), which hypothesizes that early influences multiply competencies in one domain "spill over" or are transmitted to other domains across time, culminating in positive long-term outcomes.

Three levels of system are involved in the CPC P–3 program: child, family, and school. Characteristics at each level can be influenced by the CPC program. Then those characteristics interacting with other characteristics at different levels can function as risk or protective factors of developmental outcomes. For example, the CPC program might have different effects on child's cognitive and social competence (child level), parental involvement (family level), and school mobility (school level) depending on the dosage, and the CPC effects might multiply when the effects at different domains/levels interact. Understanding the dosage effects of the CPC program on school outcomes can lead to more effective prevention programs by catalyzing cumulative advantages or "cascades" that promote positive development.

1.3.3 Contribution of the present study.—This study is unique in several respects. First, the CPC program is the only large-scale public program that reported long-term effects into adulthood (Reynolds, Ou, & Temple, 2018; Reynolds, Temple, Ou, et al., 2011), and the CPC program is still implemented nowadays. Re-examining the CPC effects in a P-3 dosage perspective provides a new look at the effectiveness of the CPC program, and it also provides insights on how gains from early childhood experiences are likely to be sustained. Second, social-emotional functioning in school is examined. Understanding the social and emotional behavior of children is the key to understand behaviors into adulthood (Jones, Greenberg, & Crowley, 2015). Third, differential effects by gender are examined. Previous studies found significant interaction effects between gender and CPC preschool participation in educational attainment favoring boys (Ou & Reynolds, 2006). However, whether dosage effects vary by gender have not been examined. Additional studies on gender differential effects will shed light on the potential pathways of the long-term effects of the CPC program. Fourth, the present study uses data from a cohort of over 1,400 participants with substantial variability in duration of participation from ages 3 to 9. Most previous studies

have small sample sizes and were not sufficiently powered to assess dosage effects. Finally, the study sample is a cohort of low-income minority children from high-poverty contexts. Findings provide insights into narrowing achievement gaps by family income and race/ethnicity and promoting healthy development.

2. Method

2.1 Sample and Design

The data were drawn from the Chicago Longitudinal Study (CLS), an on-going investigation of the well-being of a low-income minority cohort of 1,539 children who attended kindergarten in 1985–1986. Children were born in 1979 or 1980 (Reynolds, 2000). The original sample consists of 989 children who completed preschool and kindergarten in 20 CPCs and 550 children who did not attend a CPC preschool but participated in a full-day Title I kindergarten program in five randomly selected schools in impoverished Chicago areas. Title I is the largest federally funded program that provides financial assistance to schools with a high percentage of children from low-income families. Children served by the five schools matched the poverty characteristics of the CPCs.

The study sample included 1,433 participants with valid data on one or more outcomes. This is a 93.1% rate of retention. CPC participants had a higher rate of retention than the comparison group (94.1% vs. 91.3%; $p = .036$). The follow-up samples, however, well-represented the original cohort. The study used a quasi-experimental design. Like most other studies of established programs, random assignment to the intervention was not possible and would have been inappropriate given that it would have violated the legal rules requiring enrollment of the neediest children on a first-come-first-serve basis. Children in the comparison group did not enroll in the CPCs primarily because they did not live in a neighborhood containing such a center. The comparability between the program and comparison groups is well documented. Variables from many dimensions were used to account for potential selection bias (Reynolds, Temple, Ou, et al., 2011). Data were collected from participants, parents, teachers, and schools from birth to adulthood through surveys, assessments, and a variety of administrative records.

2.2 Intervention

CPC dosage was assessed by six groups based on preschool, kindergarten, and school-age participation: 1) no intervention or kindergarten only (None or K only; 9.8% had K only), 2) kindergarten and/or school-age participation (K–3; 18.7% participated between 1st and 3rd grade), 3) preschool and kindergarten participation (P–K, reference group), 4) preschool through 1st grade participation (P–1), 5) preschool through 2nd grade participation (P–2), and 6) preschool through 3rd grade participation (P–3). The differences among the six groups define the duration of participation, which allow us to examine the difference among the various combination of dosage including the group did not participate in preschool but participated in kindergarten or school-age program. The P-K group was the reference group given the policy implications in whether preschool effects fade out after the transition to elementary school. Moreover, the differences between the P-K group and other groups (no intervention or K only and K–3) who did not participate in preschool have implications in

whether school quality alone might help to narrow the achievement gaps. Table 1 presents the preprogram characteristics of intervention groups.

2.3 Outcomes

To respond to the argument on persistent and fadeout effects right after preschool participation, we selected outcomes measured between ages 9 (3rd grade) and 16 (10th grade), and outcomes that predict adult well-being (Herbers, Reynolds, & Chen, 2013; Ou, Mersky, Reynolds, & Kohler, 2007; Ou & Reynolds, 2008; Ou & Reynolds, 2010a; Ou & Reynolds, 2010b). The outcomes were classified into three categories, a) academic functioning, b) social-emotional functioning in school, and c) school experiences, to represent school-age children's well-being.

2.3.1 Academic functioning.—

Academic functioning was measured by standardized test scores in reading comprehension and math achievement on the Iowa Tests of Basic Skills (ITBS) (Hoover et al., 1993). ITBS has been tested with over 100,000 students to demonstrate high reliability and predictive validity. Reading and math achievement in the 3rd grade (ages 9–10, α= .91 and .94) and 8th grade years (ages 14–15, α = .92 and .95) were examined.

2.3.2 Social-emotional functioning in school.—

Four measures of social-emotional development were assessed: classroom adjustment in grade 3 (ages 9–10) and in grades 4–6 (ages 10–13), perceived competence in grades 3–4 (ages 9–11) and in grades 5–6 (ages 11–13). Classroom adjustment was measured on a 6-item scale (Social-Emotional Maturity Scale, SEMAT) rated by teachers from grades 3 through 6 (ages 9 to 13) (Reynolds, 2000). The six items are "child concentrates on work,", "child follows directions," "child is self-confident," "child participates in group discussions," "child gets along with others," and "child takes responsibility for his/her actions". These items were coded on a Likert-type scale from 1 (poor/not at all) to 5 (excellent/very much). The reliabilities are 0.91, 0.91, 0.89, and 0.91 for grades 3, 4, 5 and 6, respectively. The scores at grade 3 and the average scores in grades 4–6 were used (range: 6 to 30).

Perceived competence was measured by a student-rated self-concept of task persistence scale (Reynolds, 2000). Examples of the items include "my classmates like me," "I get along well with others," "I am smart," and "I try hard in school." The items were rated from 1 (disagree) to 3 (agree) in grades 3 (ages 9–10) and 4 (ages 10–11). The scale has 10 items in grade 3 but 12 items in grade 4. We prorated grade 3 scores of the 10 items into the 12-item scale for grade 4. For example, a score of 25 from 10-item in grade 3 becomes 30 when prorated into the grade 4 scale (25 × 12/10). We averaged scores between grades 3 and 4. Scores ranged from 14 to 36. The items were rated from 1 (strongly disagree) to 4 (strongly agree) in grades 5 (ages 11–12) and 6 (ages 12–13). The scale has 11 items in grade 5 but 12 items in grade 6. We prorated grade 5 scores of the 11 items into the 12-item scale for grade 6. The average scores ranged from 22 to 48. The internal-consistency reliabilities are 0.69, 0.76, 0.71, and 0.74 for grades 3, 4, 5 and 6, respectively.

2.3.3 School Experiences.—Four measures were included: grade retention, years of special education, school mobility, and expectations to go to college. Grade retention was a dichotomous measure. Children were coded 1 if they were ever retained from grades 1 through 7 (ages 6–15). Children were coded 0 if they were on grade level (grade 8) at the beginning of the 1993–94 school year. Special education is measured by a continuous variable indicating years receiving special education services. School mobility is a dichotomous measure indicating whether one has changed schools between grades 4 and 8 (ages 10–15). While reasons for school transitions might vary, the lack of continuity of school instruction with teachers and peers might disrupt the quality of school experiences. Educational expectations were measured by a dichotomous variable indicating whether the student expects to attend college or not. The survey question was asked in both grades 4 and 10: "How far in school do you think you will get?" The response in 10^{th} grade was the primary indicator with the earlier response used if missing.

2.4 Covariates

Many preprogram characteristics were included as covariates, including gender, race/ethnicity, age, maternal education and employment, marital status, teen parenthood status at child's birth, number of children in the household, participation in the Aid to Families with Dependent Children (AFDC) program, eligibility for subsidized meals, and neighborhood poverty. These were measured from birth to age 3. These indicators were also included as input variables assessing potential attrition and selection bias described below. Multiple imputation using the Expectation-Maximization procedure was used to impute missing values after determining that data were missing at random. Table 2 shows the descriptive statistics of outcomes by intervention groups.

2.5 Statistical Analysis

Several steps were taken to correct potential biases resulting from selection and attrition. First, preprogram characteristics were included in the models as covariates to adjust for any differences in child, family, and neighborhood characteristics. Second, inverse probability weighting (IPW), a propensity score technique, was used to adjust for possible biases in program selection and nonrandom attrition (Huber, 2014; Imbens & Wooldridge, 2009; Linden & Adams, 2010; Seaman & White, 2013; Wooldridge, 2007). Attrition rates vary by outcomes and intervention groups (See Table 2). Over 20 explanatory variables were used to construct the weights on attrition (See Appendix B) and program selection (See Appendix C). Research on IPW shows that it yields the most efficient estimator and is robust to misspecification (Stuart, 2010; Wooldridge, 2002).

Here is an example of how we construct the weights to correct for attrition. We first predict the probability of being in the retained sample (R=1; otherwise=0) using a probit regression model as follows:

$$R = \beta X_i + \mu_i$$

$$and\ estimates: \hat{\rho}_i = Pr(R = 1 \mid X)$$

Where R indicates whether the individual is in the retained sample or not, X is a vector of predictors that explain sample recovery and μ is an error term. The regression then estimatesρ_I, the probability of sample recovery. Then, the inverse of this predicted probability (1 /ρ) was used as a weight variable in all outcome models after verifying that our rich array of Xs variables predict sample retention and that propensity distributions between groups overlapped. Our main outcome model is:

$$Y_i = \beta_0 + \beta_1 PK3 + \beta_2 PK2 + \beta_3 PK1 + \beta_4 noPKbutSchool + \beta_5 noCPC + X_i\gamma + \varepsilon_i$$

Where Y represents our outcome variable for each participant i.

Third, we corrected for multiple comparison groups and utilized modified significance levels (Benjamini & Hochberg, 1995; Newson & Team, 2003). This procedure reduces the likelihood of mistakenly concluding that the differences in means among groups are statistically different from zero (Benjamini & Hochberg, 1995). Finally, standard errors were adjusted for site clustering, although previous studies found that corrections for clustering and alternative covariate specifications did not affect estimates (Reynolds, Temple, Ou, et al., 2011).

To evaluate the differential effects by gender, the same procedures were conducted separately for boys and girls. Previous studies have found that examining boys and girls together ignores potentially large differences in treatment effects (García et al., 2018). Separate regressions for boys and girls provide the most complete information and do not have the assumption that the associations between covariates and outcomes are the same for each gender.

Dichotomous and continuous variables were analyzed with probit and ordinary least squares regressions, respectively. One count variable (years of special education) was analyzed with zero-inflated negative binomial regression given the excessive number of zeros in the outcome variable. Regression models included the four steps described previously. Coefficients were transformed into marginal effects. Adjusted means/rates for the outcomes were calculated using the marginal effects. The P-K group was the reference group. The other five dummy variables for the CPC dosage were included in the models. These contrasts assessed whether children receiving different CPC dosage were associated with youth outcomes than those who received the P-K dosage. Effect sizes (ES) in standard deviations were calculated using Cohen's d (Cohen, 1988). Values of 0.20 and above in absolute value were interpreted as practically significant. Statistical power calculation was conducted to verify if the sample sizes were large enough to detect effects with 0.80 statistical power at the .05 level of significance (See Appendix D). Data were analyzed in STATA 15 (StataCorp, 2017).

3. Results

Regression findings are reported as adjusted means or rates by intervention groups. Note that the focus is on the group differences relative to the P-K group. The means/rates were adjusted for preprogram characteristics, both selection and attrition by IPW, and multiple comparison groups. Standard errors were clustered at the kindergarten site level.

3.1 Dosage effects

As shown in Table 3, the pattern of findings shows a dosage-response relation—youth outcomes are enhanced as years of CPC intervention increases after receiving preschool treatment. Four results were evident. First, in academic functioning, the P–3 group had the highest scores in both grades 3 and 8 compared to the P-K group ($p < .01$). ESs were relatively large in grade 3 (.70–.74) and remained practically significant in grade 8 (.31–.33). The P–2 group had significantly higher scores in grade 3 ($p < .05$; ES = .32) relative to the P-K group. There was no significant difference in the outcomes between the P-K group and other groups (P–1, K–3 and None or K only). Second, in social-emotional functioning in school, the P-3 group had significantly higher scores on classroom adjustment at grade 3 ($p < .05$; ES = .39). There was no significant difference in other outcomes between the P-K group and other groups (P–2, P–1, K–3 and None or K only).

Third, in school experiences, the P-3 group was significantly associated with lower rates of grade retention (ES = −71, $p < .001$) and school mobility (ES = −.44, $p < .01$), and fewer years in special education (ES = −.23, $p < .01$) relative to the P-K group. The P-2 group was significantly associated with a lower rate of grade retention (ES = −.22, $p < .01$) and fewer years in special education (ES = −.26, $p < .05$) relative to the P-K group. There was no significant difference in other outcomes between the P-K group and other groups (P-2, P-1, K-3 and None or K only). Finally, additional comparisons were examined between the P-3 and P-2 groups. The P-3 group had significantly higher math and reading achievement in grade 3 and a lower rate of grade retention relative to the P-2 group. See Appendix E for effect sizes. Additional analyses were conducted using different groups as the reference group. The results are noted in Table 3 for reference only. Additional tables are available upon request.

3.2 Differences by Gender

Separate analyses were conducted for boys and girls. Similar to the total sample, a dosage-response pattern of results occurred with improved performance being a function of increasing dosage after receiving preschool treatment. The patterns, however, varied by outcomes.

3.2.1 Boys.—Table 4 presents the findings for boys. In academic functioning, the P-3 group had significantly higher scores on both math and reading achievement in grade 3 relative to the P-K group ($p < .01$; ES = .57 and .64). The K-3 group had significantly lower scores on reading in grade 8 relative to the P-K group ($p < .01$; ES = −.29). In social-emotional functioning in school, the P-3 group had significantly higher scores on classroom adjustment in grades 3 and 4–6 ($p < .001$; ES= .45 and .40) and perceived competence in

grades 3–4 and 5–6 ($p < .01$; ES =. 48 and .48) relative to the P-K group. The P-2 group had a higher score on perceived competence in grades 3–4 relative to the P-K group ($p < .05$; ES =. 32). In school experiences, the P-3 group had a significantly lower rate of grade retention ($p < .001$; ES = −.68), and fewer years in special education ($p < .001$; ES = −.42) relative to the P-K group. The P-2 group had fewer years in special education ($p < .001$; ES = −.47) relative to the P-K group. There was no significant difference in other outcomes between the P-K group and other groups (P–3, P–2, P–1, K–3 and None or K only).

3.2.2 Girls.—Table 5 presents the findings for girls. In academic functioning, the P-3 group had significantly higher scores on both math and reading achievement in grade 3 relative to the P-K group ($p < .001$, ES = .86 and .83). The P-2 group had higher scores on math and reading achievement in grade 3 relative to the P-K group ($p < .05$, ES = .37 and .35). The K-3 group had significantly higher scores on math achievement in grade 3 relative to the P-K group ($p < .01$; ES= .39). There is no significant difference found between the P-K group and other groups (P–3, P–2, P– 1, K–3 and None or K only) in any outcome of social-emotional functioning in school. In school experiences, the P-3 group had significantly lower rates of grade retention ($p < .001$; ES = −.59) and school mobility ($p < .01$; ES = −.46) and a higher rate of expectations of college attendance ($p < .01$; ES = .34) relative to the P-K group. The K-3 group had fewer years in special education ($p < .01$; ES = −.55) relative to the P-K group. The None or K only group had a significantly higher rate of school mobility relative to the P-K group ($p < .05$; ES = .29). There was no significant difference in other outcomes between the P-K group and other groups (P–3, P–2, P–1, K–3 and None or K only).

3.3 Robustness

To assess the robustness of estimates, three additional model specifications were tested: adjustments on preprogram characteristics alone, adjustments on preprogram characteristics and attrition by IPW, and adjustments on preprogram characteristics and program selection by IPW. Results remained consistent. Results also were unchanged after including word analysis scores in kindergarten as a proxy for children's cognitive skills. Additional tables are available upon request.

4. Discussion

The present study investigated the effects of CPC duration on school outcomes and whether the effects differ by gender. Several points from the findings are discussed below.

4.1 Dosage Effects of the P-3 versus P-K in Academic Functioning

Research on the dosage effects of ECE has been limited despite the policy implications of such work. Relative to the P-K group, both the P–3 and P–2 groups were significantly associated with better achievement in grade 3, but only the P–3 group was significantly associated with better achievement in grade 8. The P-K group was not significantly different from other groups (P–1, K– 3, and no intervention or K-only groups). When the P-2 duration was used as the reference group, the P-3 duration remained significantly associated with better academic functioning in grade 3. In other words, it is not only essential to participate

in the CPC preschool, but also participate in the school-age program for the gain to sustain. The results suggest that the gain in preschool in academic functioning is more likely to last into 3rd grade if preschool and kindergarten were followed by at least 2 years of a school-age program. It is worth noting that school-age program was not significantly associated with outcomes when it was examined with the preschool program as two separate components in previous CLS studies (Reynolds, Temple, Ou, et al., 2011; Reynolds et al., 2007). At a glance, it might seem that the preschool program is associated with long-term outcomes while the school-age program is not, but actually, the effects of the preschool need to be accompanied by the school-age program to yield the largest gains into elementary school or beyond. Nevertheless, as a result of the large effect size at kindergarten entry (Arteaga et al., 2014; Reynolds, 1995), preschool had independent, sustained gains over time. The findings of the present study supports, in part, the sustaining environments process as discussed by Bailey et al. (2017), that identified the later supporting environments of the children who participate that can be expected to sustain beneficial program impacts.

Moreover, our results aligned well with hypotheses and findings in the literature. For example, several studies postulate the hypothesis of dynamic complementarity (Cunha & Heckman, 2009; Heckman, 2006). In these studies, it is argued that if a parent (and/or the government through policies) invest in a child in preschool and also during the school years (complementarity), the impact of these investments will be larger than if these investments are only made during the preschool years (Johnson & Jackson, 2017). Additionally, using data from Project Star, Chetty et al. (2011) found that children who were randomly assigned to higher quality classrooms in grades K through 3 had significant and sizeable long-term outcomes (e.g. higher earnings, more savings, more likely to attend to college). Our findings also support recent initiatives like the one by the federal Race to the Top Early Learning Challenge grants that sought to create “preschool through third grade approaches to sustain improved early learning outcomes through the early elementary grades (Priority 4)” in their last round of funding (U.S. Department of Education and U.S. Department of Health and Human Services, 2011). At the state level, the legislature also seems to recognize the early school years as a sensitive period. For example, Arizona’s HB 2083 (“Kindergarten; survey; report.,” 2019), proposed in the first session of 2019, petitions for collecting annual data statewide of reading proficiency levels of K-3.

4.2 Dosage Effects of the P-3 versus P-K in School Experiences

Relative to the P-K group, both the P–3 and P–2 groups were significantly associated with lower rates of grade retention and few years in special education, but only the P-3 group was significantly associated with a lower rate of school mobility. The first is on the effects of the P–3 and P–2 duration in grade retention relative to the P-K duration. The retention rates of the P–3 and P–2 groups are 3% and 23%, respectively, relative to 33% for the P-K group. The P-K group is not significantly different from other groups (P–1, K–3 and no intervention or K only). Based on the developmental cascades theory, the magnitude of the effects of the P-3 duration relative to the P-2 duration indicates that the effects of early skills associated with the additional one year might cascade over time to influence later skills. The detrimental effects of grade retention in academic performance and social adjustment have been reported in the literature (Jimerson, 2001; Ou & Reynolds, 2010b; Pagani, Tremblay,

Vitaro, Boulerice, & McDuff, 2001). Moreover, previous CLS studies have found that one of the mechanisms explain the long-term effects of the CPC program is through reducing the rates of grade retention (Ou, 2005; Reynolds & Ou, 2011; Reynolds et al., 2004). This mechanism through reducing the rates of grade retention also supports the foot-in-the-door process discussed by Bailey et al. (2017). Both the P-3 and P-2 duration are likely to be associated with other longer-term outcomes indirectly by reducing the rates of grade retention.

The second is on the effects of the P-3 and P-2 duration in years of special education relative to the P-K duration. Special education placement is an expensive practice in schools, so the effects of the ECE program in special education is typically linked to cost savings of investment. For example, reducing the need for special education services via ECE intervention is the major outcome expected in the Pay for Success (PFS) implemented in Illinois and Utah recently to expand promising ECE programs (Temple & Reynolds, 2015). PFS is a new financing method that allows state or local governments to expand cost-effective social or education services through contributions from private investors.

A final point is on the effects of the P-3 duration in school mobility relative to the P-K duration. School mobility is linked to various negative developmental outcomes (Gruman, Harachi, Abbott, Catalano, & Fleming, 2008; Herbers et al., 2013). Moreover, school mobility was found as one of the initiators of the effect of CPC preschool (Ou, 2005; Reynolds & Ou, 2011; Reynolds et al., 2004; Reynolds, Ou, et al., 2017). Those initiators can lead to long-term outcomes through other intervening factors, such as high school graduation. One explanation on the association between the P-3 duration and low school mobility might be because parents prefer to let children stay in the neighborhood due to years of positive experience in the CPC centers or any positive experiences regarding the elementary schools that they might have heard about from the CPC centers. Not only was preschool participation associated with fewer school moves, but greater parent involvement was found to be associated with fewer school moves as well (Ou, 2005; Reynolds & Ou, 2011; Reynolds et al., 2004). In addition to parent involvement, other components of the CPC program, such as outreach services, might play a strong role in accounting for this association.

4.3 Gender Analyses

Our findings suggest that boys are driving the P–3 effects in social-emotional functioning. Gender differential effects by domains are found in other studies (García et al., 2018). The direction of our findings in gender difference matches with the concept of boys at risk (Schore, 2017). Because boys are at higher risk than girls due to slower development, experiencing longer duration of early educational enrichment is likely to benefit boys more by improving social- emotional functioning and a lower need for school remedial services. Previous studies show evidence that boys do not get as much stimulation as girls in their homes (Bertrand & Pan, 2013). Thus, improving boys' social and emotional skills might have higher payoffs in prevention programs for disruptive behavior in the classroom and ultimately for preventing later involvement in the criminal justice system. Moreover, the results suggest that the pathways of long-term benefits associated with the P-3 dosage might

vary by gender. Academic functioning, social- emotional functioning, and school experiences are all predictors of adult well-being. For example, social-emotional functioning is a predictor of crime (Jones et al., 2015). Improving those outcomes might lead to better well-being in adulthood. The findings suggest that ECE is more likely to lead to long-term benefits in adulthood via improving social-emotional functioning for boys.

It is worth noting that the K–3 group had better math achievement at third grade and fewer years in special education relative to the P-K group for girls. The findings suggest that the school- age program might benefit girls even without participating in preschool. It should be further explored whether certain components of the school-age program, such as small class size or parent involvement, have a larger impact on girls than on boys. In addition, the None or K only group had a significantly higher rate of school mobility relative to the P-K group ($p < .05$; ES= .29) while there is no significant difference between the P-K group and other groups (K–3, P–1, and P–2) for girls. This suggests that participation of either CPC preschool or school-age program is associated with a lower rate of school mobility for girls. As discussed earlier, the CPC program might be related to a lower rate of school mobility through increasing parent involvement or positive school experiences. It is possible that the dosage of the CPC has a different effect in reducing school mobility for girls. This warrants further investigation. While there is evidence of different patterns of findings for boys versus girls, when we compared the estimates of boys to those of girls, only a few differences were statistically significant (See Appendix F). Thus, the results should be taken with caution.

4.4 Strengths and Limitations

The major strengths of the present study include the prospective longitudinal cohort design and large sample size to evaluate an established P–3 program. Studies on large-scale ECE have focused on the effects of preschool, and outcomes beyond third grade are rarely examined because most of the time the effects fade out by first grade. The study sample of low-income minority children is a primary focus of prevention efforts to reduce achievement gaps and improve health over the life course.

Several limitations are also notable. First, the study utilized a quasi-experimental design, which is more challenging to inferences of effects compared to well-executed randomized experiments. However, analyses using propensity score methods for both selection and attrition yielded an interpretable and consistent pattern of findings. Second, the present study did not distinguish all different duration of the intervention. Our metric of years of intervention is a global measure and more specific indicators of length (e.g., days), frequency, and intensity (instructional hours) would likely provide greater precision. We also did not adjust for average attendance rates over time. Moreover, preschool dosage (1 vs. 2 years) and the length of the kindergarten day (part- vs. full-day) were not investigated.

Third, participants attended ECE in the 1980s, and the policy context today is different. Whether findings are generalizable to current practice is an important question, and further research is needed. For example, participation in early childhood programs is substantially higher today, as nearly 4 in 5 young children are enrolled in center-based early education (U.S. Department of Education, 2017). Such greater access may underestimate the value-added effects of innovative programs like CPC P–3 because it would be harder to find

matched comparison groups. Alternatively, there has been little progress in scaling longer-duration P-3 programs.

Consequently, the educational practices implemented in our study remain largely in place today. Initial evidence of a scale-up of the CPC P-3 program shows similar benefits in school readiness skills and parent involvement than in the CLS (Reynolds, Hayakawa, et al., 2017). Whether ECE is associated with long-term outcomes and how to sustain gains remains a critical issue even though the early childhood landscape has changed since the CLS participants attended the CPC between 1983 and 1989. Findings from the present study demonstrate that longer-duration programs can have beneficial effects. Fourth, sample sizes were relatively small when the intervention effects were examined by gender. However, IPW methods can yield valid estimates with low prevalence rates and sample sizes as low as 40 (Pirracchio, Resche-Rigon, & Chevret, 2012). Power calculation was also conducted. Finally, generalizability and reproducibility should be further investigated as well as the extent to which findings are generalizable beyond economically disadvantaged samples.

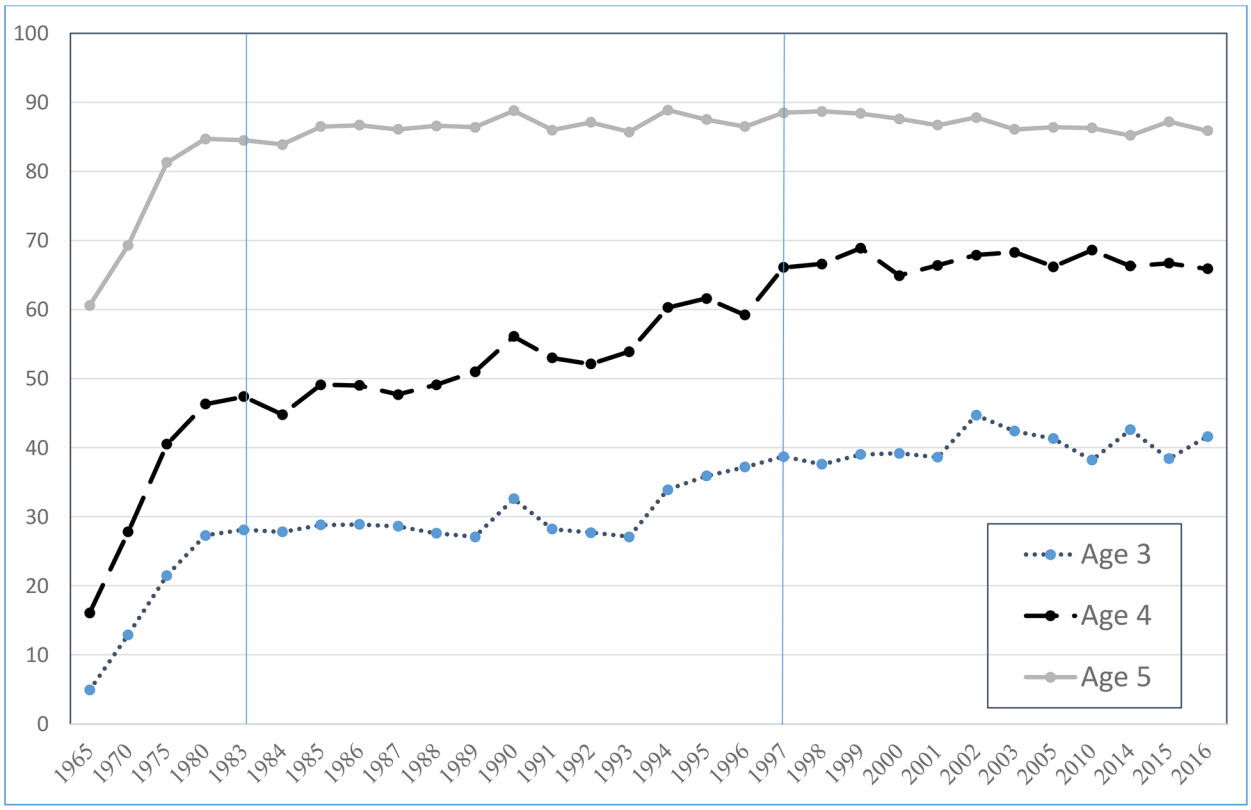
5. Conclusion

Implemented almost 30 years apart from the CPC P–3 program, findings from recent state preschool programs have shown that effects drop off after the transition to elementary school, and even if they persist (e.g.,Barnett et al., 2013; Gormley et al., 2018; Schweinhart et al., 2012), achievement gains are not as large as those found in many earlier studies of the highest quality programs (McCoy et al., 2017; Meloy et al., 2019). Our findings suggest that the P-3 dosage is associated with better academic functioning at both third and eighth grades, social-emotional functioning at third grade, and better school experiences beyond third grade compared to the P-K intervention. Preschool programs with subsequent school-age services up to third grade may be more likely to connect to better future wellness via improving academic functioning and school experiences. The P-3 dosage has the potential to further close achievement gaps by income and race/ethnicity above and beyond preschool intervention. Our findings suggest that boys are driving the P-3 effects in social-emotional functioning. Improving boys' social and emotional skills might have higher payoffs in prevention programs for disruptive behavior in the classroom. Replication and extension of findings to other locations and samples will further strengthen confidence in the diverse benefits of prevention programs for young children.

Funding:

This study was supported by the National Institute of Child Health and Human Development (Grant no. R01HD034294).

Appendix A.



3-, 4- and 5-year-old children enrolled in Early Childhood Programs Source: National Center for Education Statistics (2004, 2017)

Appendix B.

Estimates of the Marginal Effects of Being in the Retained Sample for Selected Outcomes

	3rd-grade sample	Expect. to go to college sample
Any CPC preschool	−0.007	0.039
Any CPC school age	0.045 *	0.021
Mother did not complete HS, child age 0–3	−0.014	−0.051
Child eligible for subsidized meals, child age 0–3	0.088 ***	0.041
Mother under age 18 at childbirth	−0.009	−0.081 *
Four or more children in the family, child age 0–3	0.006	−0.035
Participate in AFDC program, child age 0–3	0.056 +	−0.043 +
Mother not employed, child age 0–3	0.016	0.060 *
Single parent family status, child age 0–3	−0.014	−0.009
Indicator for missing risk factors, child age 0–3	−0.268 ***	−0.275 ***
Reside in a high poverty neighborhood	−0.024	−0.039
Low birth weight (<2500g)	0.021	−0.013
Family conflict, child age 0–5	0.054	0.079
Family financial problems, child age 0–5	−0.105 *	−0.037

	3rd-grade sample	Expect. to go to college sample
Substance abuse parent, child age 0–5	0.042	−0.097*
Female child	−0.005	0.110***
African American child	0.019	0.003
Early residential mobility	−0.063***	−0.044*
Have an SSN	0.379***	0.382***
Percent Living one year in current unit, Neighborhood census	−0.321	−0.551
Percent living 1–5 years in current unit, Neighborhood census	−0.582**	−0.481
Percent living 5–10 years in current unit, Neighborhood census	−0.366	−0.247
Percent living 10–20 years in current unit, Neighborhood census	−0.542*	−0.451
Percent Self-employed, neighborhood census	−0.120	−0.811
Percent Black families with female householder, neighborhood census	−0.025	0.025

Note:
*** p < .001,
** p < .01
* p< .05
+ p < .10.

Standard errors are clustered at the site level.

We estimate the probability of being in the retained sample using a probit regression model, controlling for child characteristics, family risk factors, neighborhood variables and having permission to track them using their social security number, and report the marginal effects in additional tables. Each column reports results for different regressions. For example, the first column reports the determinants of being in the third grade sample. Attending any CPC school-age program, having a subsidized meal and having a SSN increases the likelihood of being in the retained math sample for third grade, while family conflict and neighborhood mobility decreases the likelihood of being found in the retained sample.

Appendix C.

Estimates of Participation in P-3 Groups (n=1,531)

	P-3	P-2	P-1	P-K	K-3	None or K only
Mother did not complete HS, child age 0–3	−0.355***	−0.060	0.073	b	0.165	0.229+
Child eligible for subsidized meals, child age 0–3	0.300+	−0.387*	−0.133	b	−0.122	−0.242
Mother under age 18 at child birth	−0.053	−0.024	0.058	b	0.135	−0.040
Four or more children in family, child age 0–3	−0.114	0.076	−0.109	b	0.096	0.103
Participate in AFDC program, child age 0–3	0.078	0.266*	−0.025	b	0.224	0.091
Mother not employed, child age 0–3	−0.225	−0.174	−0.154	b	−0.177	−0.400***
Single parent family status, child age 0–3	−0.028	0.067	0.121	b	−0.126	0.050
Indicator for missing risk factors, child age 0–3	−0.747***	−0.588***	−0.113	b	−0.276	−0.079
Reside in high poverty neighborhood	−0.727	0.255	−0.203	b	−0.330	−0.225
Low birth weight (<2500g)	0.159	−0.067	0.167	b	0.236	0.146
Family conflict, child age 0–5	0.607**	0.228	0.491	b	0.258	0.413+

	P-3	P-2	P-1	P-K	K-3	None or K only
Family financial problems, child age 0–5	−0.435*	−0.364*	0.178	b	−0.034	−0.312+
Substance abuse parent, child age 0–5	−0.172	0.072	−0.418	b	−0.489*	−0.039
Female child	0.095	0.054	0.039	b	−0.189*	0.000
African American child	0.106	−0.245	−0.199	b	0.644+	−0.102
Child underage at preschool entry	−0.143	0.277	0.102	b	−0.931	−0.489
Age in months at kindergarten	0.060**	0.038*	0.032*	b	0.042**	0.039*
Constant term	−3.473	−2.166*	−2.254*	b	−2.374+	−1.331

Note: b= Base category: P-K.

*** p < .001,

** p < .01

* p< .05

+ p < .10.

Standard errors are clustered at the site level.

This table reports differences in observed characteristics according to CPC duration. The probability of receiving none or K only, K up to 3 or 1 up to 3, P–K, P–1, P–2, and P–3 is estimated using a multinomial probit regression, controlling for child characteristics and family risk factors. Overall, groups were very similar on most background variables including mother's age at childbirth, number of children in the household, participation in the AFDC program, single-parent family, low birth-weight child, family resides in a high poverty neighborhood, being an African American child, and child under age at preschool entry. Although children who attended P-3 seemed a little more advantaged than the P-K group (mothers are less likely to not complete high school, children are slightly likely to be a few months older, families are less likely to have financial problems), they also live in more conflictive families. We found similar results when we compare the P–2 group with the P-K group. However, an interesting finding is that we also found that children who did not attend any CPC program were slightly older and had mothers who were less likely to be unemployed than the P-K group. Because this might be an indicative of potential selection bias, we corrected for it using an inverse probability weighting approach. The estimated probabilities generated by the multinomial probit regression are the propensity scores used to construct the weight. The estimation assigns each individual a predicted probability of receiving each of the treatment levels.

Appendix D.

Power Calculation

	Overall	Gender
	minimum detectable effect	minimum detectable effect
Math achievement, grade 3	0.096	0.136
Math achievement, grade 8	0.098	0.136
Reading achievement, grade 3	0.100	0.136
Reading achievement, grade 8	0.098	0.136
Classroom adjustment, grade 3	0.099	0.141
Classroom adjustment, grades 4–6	0.103	0.140
Perceived competence, grades 3–4	0.099	0.141
Perceived competence, grades 5–6	0.111	0.155
Ever retained, grades 1–8, %	0.076	0.106
Years in special education, grade 1-	0.096	0.136
Ever school mobility, grades 4–8, %	0.083	0.121
Expectations of college attendance, grade 10, %	0.088	0.126

Note: Minimum detectable effect (MDE) shows that with a given sample size what is the minimum effect we can expect to measure with certain sensitivity and specificity. We used a two-tailed test, and statistical power=0.80, and alpha=0.05 (significance), as it is typically used in the literature. We used STATA 15.0 to conduct the MDE calculations. This means, for example, that if "ever retained" drops from 27% to at least 19.4% (in 7.6%=MDE) for the overall sample, our models will be able to detect the effect.

Appendix E.

Effect Sizes: Adjusted for Selection, Attrition, and Preprogram Characteristics

Outcomes	P-3 n=173	P-2 n=377	P-1 n=134	P-K n= 304	K-3 n=166	None or K only n=377	F-test
Academic Functioning							
Math achievement, grade 3	0.70***	0.32*	0.10	a	0.20	−0.12	14.87***
Math achievement, grade 8	0.31**	0.22	−0.01	a	−0.08	−0.13	7.95***
Reading achievement, grade 3	0.74***	0.32*	0.13	a	0.24	−0.06	10.87***
Reading achievement, grade 8	0.33***	0.18	−0.01	a	−0.12	−0.13	13.84***
Social-emotional Functioning							
Classroom adjustment, grade 3	0.39*	0.15	0.08	a	0.03	−0.01	1.70
Classroom adjustment, grades 4–6	0.26	0.10	0.06	a	−0.07	−0.06	5.61**
Perceived competence, grades 3–4	0.28	0.09	0.01	a	−0.12	−0.02	3.95**
Perceived competence, grades 5–6	0.38	0.05	−0.13	a	−0.14	−0.09	1.69
School Experience and Expectations							
Ever retained, grades 1–8, %	−0.71***	−0.22**	−0.04	a	−0.16	0.11	81.08***
Years in special education, grade 1–8	−0.23**	−0.26*	−0.01	a	0.06	0.01	4.37**
Ever school mobility, grades 4–8, %	−0.44**	−0.23	0.02	a	−0.06	0.16	30.22***
Expectations of college attendance, grade 10, %	0.17	0.06	0.03	a	−0.11	0.00	6.05

Note: Effect sizes and rates on the outcomes are adjusted for preprogram characteristics (i.e., child's gender, child's age, child's race, family variables, and socio-demographic factors), selection and attrition by IPW, and Benjamin-Hochberg correction for multiple comparisons. Standard errors are clustered at the site-level.

*** p<0.001,

** p< 0.01,

* p<0.05,

+ p<0.10.

F-test: joint significance of the 6-group coefficients. a Base category (reference group).

Appendix F.

Differences in Adjusted Means for Boys and Girls

Outcomes	P-3 n=173	P-2[1] n=377	P-1 n=134	P-K n= 304	K-3 n=166	None or K only n=377
Academic Functioning						
Math achievement, grade 3	−4.44	−2.31	−1.63	−0.43	−5.40	−5.36+
Math achievement, grade 8	−6.00*	−4.22	−7.55+	−4.21	−9.57**	−4.46

Outcomes	P-3 n=173	P-2[1] n=377	P-1 n=134	P-K n= 304	K-3 n=166	None or K only n=377
Reading achievement, grade 3	−7.85	−6.01 *	−7.42 +	−4.14	−4.67	−8.40 *
Reading achievement, grade 8	−5.07	−9.55 *	−12.45 **	−4.19	−11.75 ***	−9.48 +
Social-emotional Functioning						
Classroom adjustment, grade 3	−1.95 *	−1.84 +	−4.62 ***	−2.91 *	−3.44 **	−2.35 *
Classroom adjustment, grades 4–6	−1.84 *	−2.19 *	−2.03 **	−2.55 *	−2 .01 +	−2.79 *
Perceived competence, grades 3–4	−0.94	−0.50	−1.25	−1.98 *	−1.92 *	−0.63
Perceived competence, grades 5–6	0.24	−0.98	−0.74	−0.39	−1.07	−1.46
School Experience and Expectations						
Ever retained, grades 1–8, %	0.09	0.19 **	0.17	0.18 +	0.23 **	0.20 *
Years in special education, grades 1–8	0.08	−0.28	0.28	0.26	0.94 **	−0.31
Ever school mobility, grades 4–8, %	0.09	−0.01	0.21	0.06	0.10	−0.05
Expectations of college attendance, grade 10, %	−0.16 *	0.00	−0.05	−0.07	−0.08	−0.09

*** p<0.001,
** p< 0.01,
* p<0.05,
+ p<0.10.

We ran separate regressions for boys and girls, and then we reported the adjusted means by CPC contrast on tables 3 and 4, respectively. This table shows the gender difference in the adjusted means (adjusted means of boys – adjusted means of girls) by CPC contrast. We then used the “adjust” command in STATA to test for statistical significance.

REFERENCES

Ansari A, & Pianta RC (2018a). The role of elementary school quality in the persistence of preschool effects. Children and Youth Services Review, 86, 120–127. doi:10.1016/j.childyouth.2018.01.025

Ansari A, & Pianta RC (2018b). Variation in the long-term benefits of child care: The role of classroom quality in elementary school. Developmental Psychology, 54(10), 1854–1867. doi: 10.1037/dev0000513 [PubMed: 29620389]

Arteaga I, Humpage S, Reynolds AJ, & Temple JA (2014). One year of preschool or two: Is it important for adult outcomes? Economics of Education Review, 40, 221–237. [PubMed: 26823640]

Autor D, Figlio D, Karbownik K, Roth J, & Wasserman M (2015). Family disadvantage and the gender gap in behavioral and educational outcomes. Northwestern University, Institute for Policy Research, Working Paper Series. Retrieved from http://www.ipr.northwestern.edu/publications/papers/2015/

Bailey D, Duncan GJ, Odgers CL, & Yu W (2017). Persistence and fadeout in the impacts of child and adolescent interventions. Journal of Research on Educational Effectiveness, 10(1), 7–39. doi: 10.1080/19345747.2016.1232459 [PubMed: 29371909]

Barnett WS, Jung K, Youn M, & Frede E (2013). Abbott preschool program longitudinal effects study: Fifth grade follow-up. New Brunswick, NJ: National Institute for Early Education Research.

Bassok D, Gibbs CR, & Latham S (2018). Preschool and children’s outcomes in elementary school: Have patterns changed nationwide between 1998 and 2010? Child Development, Advance online publication. doi:10.1111/cdev.13067

Benjamini Y, & Hochberg Y (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. Journal of the Royal Statistical Society. Series B (Methodological), 57(1), 289–300.

Benner AD, Thornton A, & Crosnoe R (2016). Children’s exposure to sustainability practices during the transition from preschool into school and their learning and socioemotional development. Applied Developmental Science, 1–14. doi:10.1080/10888691.2016.1175946

Bertrand M, & Pan J (2013). The trouble with boys: Social influences and the gender gap in disruptive behavior. American economic journal: applied economics, 5(1), 32–64.

Black MM, Walker SP, Fernald LCH, Andersen CT, DiGirolamo AM, Lu C,…Grantham-McGregor S (2017). Early childhood development coming of age: science through the life course. The Lancet, 389(10064), 77–90. doi:10.1016/S0140-6736(16)31389-7

Blok H, Fukkink R, Gebhardt E, & Leseman P (2005). The relevance of delivery mode and other programme characteristics for the effectiveness of early childhood intervention. International Journal of Behavioral Development, 29(1), 35–47. doi:10.1080/01650250444000315

Bronfenbrenner U (1989). Ecological systems theory. Annals of Child Development, 6, 187–249.

Buchmann C, DiPrete TA, & McDaniel A (2008). Gender inequalities in education. Annual Review of Sociology, 34, 319–337.

Burger K (2010). How does early childhood care and education affect cognitive development? An international review of the effects of early interventions for children from different social backgrounds. Early Childhood Research Quarterly, 25(2), 140–165.

Camilli G, Vargas S, Ryan S, & Barnett WS (2010). Meta-Analysis of the Effects of Early Education Interventions on Cognitive and Social Development. Teachers College Record, 112(3), 579–620.

Cannon JS, Kilburn MR, Karoly LA, Mattox T, Muchow AN, & Buenaventura M (2017a). Decades of evidence demonstrate that early childhood programs can benefit children and provide economic returns. Retrieved from Santa Monica, CA: https://www.rand.org/pubs/research_briefs/RB9993.html.

Cannon JS, Kilburn MR, Karoly LA, Mattox T, Muchow AN, & Buenaventura M (2017b). Investing early: Taking stock of outcomes and economic returns from early childhood programs. Santa Monica, CA: RAND Corporation.

Cascio EU, & Schanzenbach DW (2013). The impacts of expanding access to high-quality preschool education Cambridge, MA: National Bureau of Economic Research.

Chaudry A, & Datta AR (2017). The current landscape for public pre-kindergarten programs In Phillips DA, Lipsey MW, Dodge KA, Haskins R, Bassok D, Burchinal MR, Duncan GJ, Dynarski M, Magnuson KA, & Weiland C (Eds.), The Current State of Scientific Knowledge on Pre-Kindergarten Effects. Washington, DC: Brookings Institute.

Chetty R, Friedman JN, Hilger N, Saez E, Schanzenbach DW, & Yagan D (2011). How does your kindergarten classroom affect your earnings? Evidence from Project Star. The Quarterly Journal of Economics, 126(4), 1593–1660. doi:10.1093/qje/qjr041 [PubMed: 22256342]

Claessens A, Engel M, & Curran FC (2014). Academic content, student learning, and the persistence of preschool Effects. American Educational Research Journal, 51(2), 403–434. doi:10.3102/0002831213513634

Cohen J (1988). Statistical power analysis for the behavioral sciences (2nd ed.). Hillsdale, N.J.:L. Erlbaum Associates.

Conti G, Heckman JJ, & Pinto R (2015). The effects of two influential early childhood interventions on health and healthy behaviors. National Bureau of Economic Research Working Paper Series, *No. 21454*. doi:10.3386/w21454

Cunha F, & Heckman JJ (2009). The economics and psychology of inequality and human development. Journal of the European Economic Association, 7(2–3), 320–364. doi:10.1162/jeea.2009.7.2-3.320 [PubMed: 20209045]

Currie J, & Thomas D (2000). School quality and the longer-term effects of Head Start. Journal of Human Resources, 35(4), 755–774. doi:10.2307/146372

Dodge KA, Bai Y, Ladd HF, & Muschkin CG (2017). Impact of North Carolina's early childhood programs and policies on educational outcomes in elementary school. Child Development, 88(3), 996–1014. doi:10.1111/cdev.12645 [PubMed: 27859011]

Duncan GJ, & Sojourner AJ (2013). Can intensive early childhood intervention programs eliminate income-based cognitive and achievement gaps? Journal of Human Resources, 48(4), 945–968. [PubMed: 25620809]

Farran DC (2000). Another decade of intervention for children who are low income or disabled: What do we know now? In Shonkoff JP & Meisels SJ (Eds.), Handbook of early childhood intervention, 2nd ed. (pp. 510–548). New York, NY: Cambridge University Press.

Fortin NM, Oreopoulos P, & Phipps S (2015). Leaving boys behind. Journal of Human Resources, 50(3), 549–579. doi:10.3368/jhr.50.3.549

García JL, Heckman JJ, & Ziff AL (2018). Gender differences in the benefits of an influential early childhood program. European Economic Review, 109, 9–22. doi:10.1016/j.euroecorev.2018.06.009 [PubMed: 30410186]

Gilliam WS, & Zigler EF (2000). A critical meta-analysis of all evaluations of state-funded preschool from 1977 to 1998: Implications for policy, service delivery and program evaluation. Early Childhood Research Quarterly, 15(4), 441–473. doi:10.1016/S0885-2006(01)00073-4

Golding P, & Fitzgerald HE (2017). Psychology of boys at risk: Indicators from 0–5. Infant Mental Health Journal, 38(1), 5–14. doi:10.1002/imhj.21621 [PubMed: 27959473]

Gormley WT, Phillips D, & Anderson S (2018). The effects of Tulsa's pre-k program on middle school student performance. Journal of Policy Analysis and Management, 37(1), 63–87. doi:10.1002/pam.22023

Gruman DH, Harachi TW, Abbott RD, Catalano RF, & Fleming CB (2008). Longitudinal effects of student mobility on three dimensions of elementary school engagement. Child Development, 79(6), 1833–1852. [PubMed: 19037953]

Heckman JJ (2006). Skill formation and the economics of investing in disadvantaged children. Science, 312(5782), 1900–1902. doi:10.1126/science.1128898 [PubMed: 16809525]

Herbers JE, Reynolds AJ, & Chen CC (2013). School mobility and developmental outcomes in young adulthood. Development and Psychopathology, 25(2), 501–515. [PubMed: 23627959]

Hill CJ, Gormley WT, & Adelstein S (2015). Do the short-term effects of a high-quality preschool program persist? Early Childhood Research Quarterly, 32(Supplement C), 60–79. doi:10.1016/j.ecresq.2014.12.005

Hoover H, Hieronymus A, Frisbie D, Dunbar S, Oberly K, Cantor N,…Qualls-Payne A (1993). Iowa Tests of Basic Skills interpretive guide for teachers and counselors, Forms K & L, Levels 9–14. Iowa City, IA: University of Iowa.

Huber M (2014). Identifying causal mechanisms (primarily) based on inverse probability weighting. Journal of Applied Econometrics, 29(6), 920–943. doi:10.1002/jae.2341

Imbens GW, & Wooldridge JM (2009). Recent developments in the econometrics of program evaluation. Journal of economic literature, 47(1), 5–86.

Jenkins JM, Watts TW, Magnuson K, Gershoff ET, Clements DH, Sarama J, & Duncan G (2018). Do high quality kindergarten and first grade classrooms mitigate preschool fadeout? Journal of Research on Educational Effectiveness, Advance online publication. doi:10.1080/19345747.2018.1441347

Jimerson SR (2001). Meta-analysis of grade retention research: Implications for practice in the 21st century. School Psychology Review, 30(3), 420–437.

Johnson RC, & Jackson CK (2017). Reducing inequality through dynamic complementarity: Evidence from Head Start and public school spending Retrieved from https://www.nber.org/papers/w23489

Jones DE, Greenberg M, & Crowley M (2015). Early social-emotional functioning and public health: The relationship between kindergarten social competence and future wellness. American Journal of Public Health, 105(11), 2283–2290. doi:10.2105/ajph.2015.302630 [PubMed: 26180975]

Jung K, Barnett WS, Hustedt E, & Francis J (2013). Longitudinal effects of the Arkansas Better Chance program: Findings from first grade through fourth grade. New Brunswick, NJ: National Institute for Early Education Research.

Karoly LA, Kilburn MR, & Cannon JS (2005). Early childhood interventions: Proven results, future promises. Santa Monica, CA: RAND Corporation.

Kindergarten; survey; report., 2083, Arizona House Bill § 15–249.14 (2019).

Ladd HF (2017). Do some groups of children benefit more than others from pre-kindergarten programs? In Phillips DA, Lipsey MW, Dodge KA, Haskins R, Bassok D, Burchinal MR, Duncan GJ, Dynarski M, Magnuson KA, & Weiland C (Eds.), The Current State of Scientific Knowledge on Pre-Kindergarten Effects (pp. 31–36). Washington, DC: Brookings Institute.

Linden A, & Adams JL (2010). Using propensity score-based weighting in the evaluation of health management programme effectiveness. Journal of Evaluation in Clinical Practice, 16(1), 175–179. doi:10.1111/j.1365-2753.2009.01219.x [PubMed: 20367829]

Lipsey MW, Farran DC, & Hofer KG (2015). A randomized control trial of the effects of a statewide voluntary prekindergarten program on children's skills and behaviors through third grade. Retrieved from Nashville, TN: http://peabody.vanderbilt.edu/research/pri/VPKthrough3rd_final_withcover.pdf

Loveless T (2015). Girls, boys, and reading In How well are American students learning? With sections on the gender gap in reading, effects of the Common Core, and student engagement (Vol. 3, pp. 8–17). Washington, DC: The Brookings Institution.

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

Magnuson KA, Meyers MK, Ruhm CJ, & Waldfogel J (2004). Inequality in preschool education and school readiness. American Educational Research Journal, 41(1), 115–157. doi:10.3102/00028312041001115

Magnuson KA, Ruhm C, & Waldfogel J (2007). The persistence of preschool effects: Do subsequent classroom experiences matter? Early Childhood Research Quarterly, 22(1), 18–38. doi:10.1016/j.ecresq.2006.10.002

Mashburn AJ, & Yelverton R (2019). Patterns of experiences across Head Start and kindergarten classrooms that promote children's development In Reynolds AJ & Temple JA (Eds.), Sustaining Early Childhood Learning Gains: Program, School, and Family Influences (pp. 135–159). New York, NY: Cambridge University Press.

Masten AS, & Cicchetti D (2010). Developmental cascades. Development and Psychopathology, 22(3), 491–495. doi:10.1017/S0954579410000222 [PubMed: 20576173]

McCoy DC, Yoshikawa H, Ziol-Guest KM, Duncan GJ, Schindler HS, Magnuson K,…Shonkoff JP (2017). Impacts of early childhood education on medium- and long-term educational outcomes. Educational Researcher, 46(8), 474–487. doi:10.3102/0013189x17737739 [PubMed: 30147124]

Meloy B, Gardner M, & Darling-Hammond L (2019). Untangling the evidence on preschool effectiveness: Insights for policymakers: Research Brief. Palo Alto, CA: Learning Policy Institute.

Moore JE, Cooper BR, Domitrovich CE, Morgan NR, Cleveland MJ, Shah H,…Greenberg MT (2015). The effects of exposure to an enhanced preschool program on the social-emotional functioning of at-risk children. Early Childhood Research Quarterly, 32, 127–138.

Muennig P (2015). Can universal pre-kindergarten programs improve population health and longevity? Mechanisms, evidence, and policy implications. Social Science & Medicine, 127, 116–123. [PubMed: 25174771]

Muschkin CG, Ladd HF, Dodge KA, & Bai Y (2018). Gender differences in the impact of North Carolina's early care and education initiatives on student outcomes in elementary school. Educational Policy, Advance online publication. doi:10.1177/0895904818773901

National Center for Education Statistics. (2004). Digest of Education Statistics 2003. Retrieved from Washington DC: U.S. Department of Education: https://nces.ed.gov/pubs2005/2005017.pdf

National Center for Education Statistics. (2017). Digest of Education Statistics 2016. Retrieved from Washington DC: U.S. Department of Education: https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2018138

National Institute of Early Education Research (NIEER). (2018). Preschool Yearbook, 2017. Retrieved from New Brunswick, NJ: Rutgers University:

Newson R, & Team AS (2003). Multiple-test procedures and smile plots. The Stata Journal, 3(2), 109–132. doi:10.1177/1536867X0300300202

Ou S-R (2005). Pathways of long-term effects of an early intervention program on educational attainment: Findings from the Chicago longitudinal study. Journal of Applied Developmental Psychology, 26(5), 578–611.

Ou S-R, Mersky JP, Reynolds AJ, & Kohler KM (2007). Alterable predictors of educational attainment, income, and crime: Findings from an inner-city cohort. Social Service Review, 81(1), 85–128.

Ou S-R, & Reynolds AJ (2006). Early childhood intervention and educational attainment: Age 22 findings from the Chicago Longitudinal Study. Journal of Education for Students Placed at Risk (JESPAR), 11(2), 175–198. doi:10.1207/s15327671espr1102_4

Ou S-R, & Reynolds AJ (2008). Predictors of educational attainment in the Chicago Longitudinal Study. School Psychology Quarterly, 23(2), 199–229.

Ou S-R, & Reynolds AJ (2010a). Childhood predictors of young adult male crime. Children and Youth Services Review, 32(8), 1097–1107. [PubMed: 20657803]

Ou S-R, & Reynolds AJ (2010b). Grade retention, postsecondary education, and public aid receipt. Educational Evaluation and Policy Analysis, 32(1), 118–139.

Pagani L, Tremblay RE, Vitaro F, Boulerice B, & McDuff P (2001). Effects of grade retention on academic performance and behavioral development. Development and Psychopathology, 13(2), 297–315. doi:10.1017/S0954579401002061 [PubMed: 11393648]

Phillips DA, Lipsey MW, Dodge KA, Haskins R, Bassok D, Burchinal MR,…Weiland C (2017). Puzzling it out: The current state of scientific knowledge on pre- kindergarten effects A consensus statement In Phillips DA, Lipsey MW, Dodge KA, Haskins R, Bassok D, Burchinal MR, Duncan GJ, Dynarski M, Magnuson KA, & Weiland C (Eds.), The current state of scientific knowledge on pre- kindergarten effects. Washington, DC: Brookings Institute.

Pirracchio R, Resche-Rigon M, & Chevret S (2012). Evaluation of the propensity score methods for estimating marginal odds ratios in case of small sample size. BMC Medical Research Methodology, 12(1), 70. doi:10.1186/1471-2288-12-70 [PubMed: 22646911]

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). Retrieved from https://www.acf.hhs.gov/sites/default/files/opre/head_start_report.pdf

Ramey CT, & Ramey SL (2019). Reframing policy and practice deliberations: Twelve hallmarks of strategies to attain and sustain early childhood gains In Reynolds AJ & Temple JA (Eds.), Sustaining Early Childhood Learning Gains: Program, School, and Family Influences (pp. 314–349). New York, NY: Cambridge University Press.

Reynolds AJ (1995). One-Year of Preschool Intervention or 2 - Does It Matter. Early Childhood Research Quarterly, 10(1), 1–31.

Reynolds AJ (2000). Success in early intervention: The Chicago Child-Parent Centers. Lincoln, NE: University of Nebraska Press.

Reynolds AJ, Hayakawa M, Candee A, & Englund MM (2016). CPC P-3 Program Manual: Child-Parent Center Preschool–3rd Grade Program. Minneapolis, MN: University of Minnesota, Human Capital Research Collaborative.

Reynolds AJ, Hayakawa M, Ou S-R, Mondi CF, Englund MM, Candee AJ, & Smerillo NE (2017). Scaling and sustaining effective early childhood programs through school–family–university collaboration. Child Development, 88(5), 1453–1465. doi:10.1111/cdev.12901 [PubMed: 28768056]

Reynolds AJ, Magnuson KA, & Ou S-R (2010). Preschool-to-third grade programs and practices: A review of research. Children and Youth Services Review, 32(8), 1121–1131. [PubMed: 27667887]

Reynolds AJ, & Ou S-R (2011). Paths of effects From preschool to adult well-being: A confirmatory analysis of the Child-Parent Center program. Child Development, 82(2), 555–582. doi:10.1111/j.1467-8624.2010.01562.x [PubMed: 21410923]

Reynolds AJ, Ou S-R, & Temple JA (2018). A multicomponent, preschool to third grade preventive intervention and educational attainment at 35 years of age. JAMA Pediatrics, 172(3), 247–256. doi:10.1001/jamapediatrics.2017.4673 [PubMed: 29379955]

Reynolds AJ, Ou S-R, & Topitzes JW (2004). Paths of effects of early childhood intervention on educational attainment and delinquency: A confirmatory analysis of the Chicago Child-Parent Centers. Child Development, 75(5), 1299–1328. [PubMed: 15369516]

Reynolds AJ, Ou SR, Mondi CF, & Hayakawa M (2017). Processes of Early Childhood Interventions to Adult Well–Being. Child Development, 88(2), 378–387. [PubMed: 28195326]

Reynolds AJ, & Temple JA (2019). Introduction: Increasing and sustaining gains in early learning In Reynolds AJ & Temple JA (Eds.), Sustaining Early Childhood Learning Gains: Program, School, and Family Influences (pp. 1–27). New York, NY: Cambridge University Press.

Reynolds AJ, Temple JA, Ou S-R, Arteaga IA, & White BAB (2011). School-based early childhood education and age-28 well-being: Effects by timing, dosage, and subgroups. Science, 333(6040), 360–364. [PubMed: 21659565]

Reynolds AJ, Temple JA, Ou S-R, Robertson DL, Mersky JP, Topitzes JW, & Niles MD (2007). Effects of a school-based, early childhood intervention on adult health and well-being - A 19-year follow-up of low-income families. Archives of Pediatrics & Adolescent Medicine, 161(8), 730–739. [PubMed: 17679653]

Reynolds AJ, Temple JA, Robertson DL, & Mann EA (2001). Long-term effects of an early childhood intervention on educational achievement and juvenile arrest - A 15-year follow-up of low-income children in public schools. Jama-Journal of the American Medical Association, 285(18), 2339–2346.

Reynolds AJ, Temple JA, White BAB, Ou S, & Robertson DL (2011). Age 26 cost- benefit analysis of the Child-Parent Center early education program. Child Development, 82(1), 379–404. [PubMed: 21291448]

Rutter M, & Rutter M (1993). Developing minds: challenge and continuity across the life span. New York, NY: BasicBooks.

Schore AN (2017). All our sons: The developmental neurobiology and neuroendocrinology of boys at risk. Infant Mental Health Journal, 38(1), 15–52. doi:10.1002/imhj.21616 [PubMed: 28042663]

Schweinhart LJ, Montie J, Xiang Z, Barnett WS, Belfield CR, & Nores M (2005). Lifetime effects: the High/Scope Perry preschool study through age 40. Ypsilanti, MI: High/Scope Press.

Schweinhart LJ, Xiang Z, Daniel-Echols M, Browning K, & Wakabayashi T (2012). Michigan Great Start Readiness Program evaluation 2012: High school graduation and grade retention findings. Ypsilanti, MI: HighScope Educational Research Foundation.

Seaman SR, & White IR (2013). Review of inverse probability weighting for dealing with missing data. Statistical Methods in Medical Research, 22(3), 278–295. doi:10.1177/0962280210395740 [PubMed: 21220355]

StataCorp. (2017). Stata: Release 15. College Station, TX: StataCorp LP.

Stuart EA (2010). Matching methods for causal inference: A review and a look forward. Statistical science: a review journal of the Institute of Mathematical Statistics, 25(1), 1–21. [PubMed: 20871802]

Swain WA, Springer MG, & Hofer KG (2015). Early grade teacher effectiveness and pre- k effect persistence: Evidence from Tennessee. AERA Open, 1(4), 1–17. doi:10.1177/2332858415612751 [PubMed: 28164148]

Temple JA, & Reynolds AJ (2015). Using social-impact borrowing to expand preschool-to- third grade programs in urban schools. Journal of Education for Students Placed at Risk (JESPAR), 20(4), 281–292. doi:10.1080/10824669.2015.1076341 [PubMed: 26834449]

U.S. Department of Education. (2015). Preschool Development Grants Program Brochure Years 1 and 2. Retrieved from http://www2.ed.gov/programs/preschooldevelopmentgrants/pdgbrochure1and2.pdf

U.S. Department of Education. (2017). Digest of education statistics, 2016. Washington, DC.

U.S. Department of Education and U.S. Department of Health and Human Services. (2011). Race to the top-Early learning challenge. Retrieved from https://www.ed.gov/early-learning/elc-draft-summary

Wooldridge JM (2002). Inverse probability weighted M-estimators for sample selection, attrition, and stratification. Portuguese Economic Journal, 1(2), 117–139.

Wooldridge JM (2007). Inverse probability weighted estimation for general missing data problems. Journal of Econometrics, 141(2), 1281–1301. doi:10.1016/j.jeconom.2007.02.002

Yoshikawa H, Weiland C, Brooks-Gunn J, Burchinal MR, Espinosa LM, Gormley WT,…Zaslow MJ (2013). Investing in our future: The evidence base on preschool education. Retrieved from http://fcd-us.org/resources/evidence-base-preschool

Zhai F, Raver CC, & Jones SM (2012). Academic performance of subsequent schools and impacts of early interventions: Evidence from a randomized controlled trial in Head Start settings. Children and Youth Services Review, 34(5), 946–954. [PubMed: 22773872]

- P-3 dosage is associated with better school outcomes than P-K, P–1, or P–2 dosage. (83)
- Performance of the P-K group is not different from no intervention, K-3 or P-1 groups. (83)
- Boys are driving the P-3 effects in social-emotional functioning. (66)
- K-3 dosage is associated with 2 outcomes positively than P-K dosage for girls. (82)
- Girls not participate in CPC are associated with higher rates of school mobility. (82)

Table 1.

Comparability of Preprogram Attributes by Intervention Groups

Variables	Study sample	P-3 (5–6 yr)	P-2 (4–5 yr)	P-1 (3–4 yr)	P-K[1] (2–3 yr)	K-3 (1–4 yr)	None/K only (0–1 yr)
Sample size	1531	173	377	134	304	166	377
Characteristics							
Mother did not complete HS, child age 0–3	54%	40% **	52%	56%	54%	61%	59%
Child eligible for subsidized meals, child age 0–3	84%	88%	80% *	84%	87%	86%	81% +
Mother under age 18 at child birth	16%	13%	15%	19%	16%	20%	16%
Four or more children in family, child age 0–3	17%	14%	18%	13%	16%	18%	18%
Participate in AFDC program, child age 0–3	63%	60%	64%	60%	65%	69%	58%
Mother not employed, child age 0–3	66%	64% +	66%	66%	71%	72%	61% **
Single parent family status, child age 0–3	76%	75%	77%	79%	76%	75%	76%
Indicator for missing risk factors, child age 0–3	16%	9% ***	10% ***	19%	22%	17%	18%
Reside in high poverty neighborhood	76%	58% ***	85% +	75%	81%	72% **	74% **
Low birth weight (<2500g)	12%	12%	10%	13%	10%	16%	12%
Child underage at preschool entry	3%	2%	1%	1%	1%	8% ***	4% *
Family conflict, child age 0–5	6%	8% +	5%	8% +	4%	5%	6%
Family financial problems, child age 0–5	7%	6%	6%	11%	8%	8%	6%
Substance abuse parent, child age 0–5	4%	4%	5%	4%	4%	2%	5%
Female child	50%	54%	53%	50%	50%	42%	49%
African American child	93%	95%	92%	91%	93%	98% *	92%
Age in months at kindergarten	63.3	63.8 ***	63.2 *	63.2	62.6	64.1 ***	63.5 **
Sample recovery rate							
% with achievement by 3rd grade	84%	99% ***	92% ***	76%	72%	83% *	82% **
% with achievement by 8th grade	87%	98% ***	94% ***	86% +	78%	89% **	84%
% with school mobility	88%	99% ***	95% ***	86%	79%	90% **	86% +
% with perceived competence, grades 3–4	80%	97% ***	90% ***	69%	66%	80% ***	76% ***
% with perceived competence, grades 5–6	61%	70% ***	62% **	61% *	52%	61% *	63% ***

Variables	Study sample	P-3 (5–6 yr)	P-2 (4–5 yr)	P-1 (3–4 yr)	P-K[1] (2–3 yr)	K-3 (1–4 yr)	None/K only (0–1 yr)
% with expectations of college attendance	72%	84% ***	79% ***	69%	66%	66%	68%

*** p < .001,

** p < .01

* p< .05

+ p < .10

The original sample of 1,539 was reduced to 1,531 due to 8 participants not having identifying information for matching.

[1] Reference group

Author Manuscript
Author Manuscript
Author Manuscript

Table 2.

Descriptive Statistics of Outcomes by Intervention Groups

Outcomes	Sample size	Overall mean	P-3	P-2	P-1	P-K	K-3	None or K only
Academic Functioning								
Math achievement, grade 3	1,289	100.66	107.58	102.41	99.78	98.27	99.95	97.18
Math achievement, grade 8	1,343	147.32	152.53	151.04	146.47	145.55	143.09	143.95
Reading achievement, grade 3	1,288	97.05	106.33	99.08	95.49	93.50	97.63	92.37
Reading achievement, grade 8	1,344	144.72	151.18	148.38	144.56	143.34	139.31	140.77
Social-emotional Functioning								
Classroom adjustment, grade 3	1,326	18.92	20.63	19.46	18.99	18.13	17.99	18.36
Classroom adjustment, grades 4–6	1,224	18.72	19.58	19.49	18.77	18.24	17.98	18.11
Perceived competence, grades 3–4	1,227	28.33	29.04	28.60	28.22	28.11	27.78	28.04
Perceived competence, grades 5–6	1,054	33.65	33.64	34.28	33.39	34.07	33.35	33.10
School Experiences and Expectation								
Ever retained, grades 1–8, %	1,377	27.7%	9.2%	22.0%	31.9%	32.8%	28.9%	37.9%
Years in special education, grades 1–8	1,377	0.60	0.34	0.30	0.73	0.67	0.95	0.82
Ever school mobility, grade 4–8, %	1,361	60.6%	39.2%	53.3%	67.0%	65.7%	63.1%	72.8%
Expectations of college attendance, grade 10, %	1,102	84.8%	89.0%	86.0%	85.9%	83.0%	80.0%	84.0%

Table 3.

Adjusted Means for Outcomes: Adjusted for Selection, Attrition, and Preprogram Characteristics

Outcomes	P-3 n=173	P-2 n=377	P-1 n=134	P-K[1] n= 304	K-3 n=166	None or K only n=377	ES range relative to P-K	F-test
Academic Functioning								
Math achievement, grade 3	107.19 ***a	102.23 *b	99.52	98.27	100.60	96.37	−.12–.70	14.87 ***
Math achievement, grade 8	152.21 **b	150.62	146.36	146.53	145.17	144.14	−. 13–.31	7.95 ***
Reading achievement, grade 3	104.93 ***a	97.85 *c	95.09	93.02	96.48	91.49	−.06–.74	10.87 ***
Reading achievement, grade 8	151.24 ***b	148.01	143.76	143.99	141.53	141.30	−. 13–.33	13.84 ***
Social-emotional Functioning								
Classroom adjustment, grade 3	20.48 *c	19.14	18.78	18.33	18.53	18.30	−.01–.39	1.70
Classroom adjustment, grades 4–6	19.72	18.96	18.80	18.52	18.06	18.16	−.07–.26	5.61 **
Perceived competence, grades 3–4	29.18	28.45	28.17	28.13	27.68	28.07	−.12–.28	3.95 **
Perceived competence, grades 5–6	35.78	34.25	33.46	33.99	33.45	33.66	−. 14–.38	1.69
School Experience and Expectations								
Ever retained, grades 1–8, %	0.03 ***a	0.23 **	0.31	0.33	0.26	0.38	−.71–11	81.08 ***
Years in special education, grade 1–8	0.87 **c	1.00 *c	1.26	1.31	1.26	1.38	−.26–.06	4.37 **
Ever school mobility, grades 4–8, %	0.44 **b	0.54	0.66	0.65	0.62	0.73	−.44–.16	30.22 ***
Expectations of college attendance, grade 10, %	0.90	0.86	0.85	0.84	0.80	0.84	−.11–17	6.05

Note: Adjusted means/rates on the outcomes are calculated using the marginal means, which were adjusted for preprogram characteristics (i.e., child's gender, child's race, child's age, family variables, and socio-demographic factors), selection and attrition by IPW, and Benjamin-Hochberg correction for multiple comparisons. Standard errors are clustered at the site level at the beginning of the study.

[1]Reference group.

F-test: joint significance of the 6-group coefficients.

*** p<0.001,

** p< 0.01,

* p<0.05,

+ p<0.10.

[a]Significant compared to four reference groups: P-2, P-1, P-K, None or K only.

[b]Significant compared to three reference groups: P-1, P-K, None or K only.

[c]Significant compared to two reference groups: P-K, None or K only.

Table 4.

Adjusted Means for Outcomes Adjusted for Selection, Attrition, and Preprogram Characteristics, Boys (N = 763)

Outcomes	P-3 n=79	P-2 n=178	P-1 n=67	P-K[1] n= 152	K-3 n=96	None or K only n=191	ES range relative to P-K	F-test
Academic Functioning								
Math achievement, grade 3	105.07 ***	100.85	99.21	98.08	98.08	93.61	−.20–.64	10.06 ***
Math achievement, grade 8	149.17	148.21	143.09	144.56	140.71	141.91	−.25–.33	2.37 +
Reading achievement, grade 3	101.04 **	94.60	91.40	91.18	95.39	87.08	−.20–.32	5.36 **
Reading achievement, grade 8	149.12	143.10	138.10	142.25	136.01 **	137.04	−.29–.32	4.48 **
Social-emotional Functioning								
Classroom adjustment, grade 3	19.52 ***	18.23	16.51	16.76	16.74	17.12	−.05–.50	4.90 **
Classroom adjustment, grade 4–6	18.99 ***	17.89	17.70	17.10	16.88	16.67	−.07–.40	6.37 ***
Perceived competence, grades 3–4	28.71 ***	28.13 *	27.51	26.98	26.77	27.75	−.04–.48	5.72 **
Perceived competence, grades 5–6	35.99 **	33.76	33.17	33.78	33.19	33.03	−. 19–.48	8.28 **
School Experiences and Expectation								
Ever retained, grades 1–8, %	0.11 ***	0.33	0.39	0.42	0.37	0.49	−.50–.14	47.82 **
Years in special education, grades 1–8	0.92 ***	0.87 ***	1.41	1.39	1.36	1.29	−.42–.17	5.93 **
Ever school mobility, grades 4–8, %	0.49	0.55	0.78	0.68	0.68	0.71	−.27–.10	11.75 *
Expectations of college attendance, grade 10, %	0.83	0.87	0.82	0.81	0.81	0.81	−. 10–. 17	3.91

Note: Adjusted means/rates on the outcomes are calculated using the marginal means, which were adjusted for preprogram characteristics (i.e., child's gender, child's race, child's age, family variables, and socio-demographic factors), selection and attrition by IPW, and Benjamin-Hochberg correction for multiple comparisons.. Standard errors are clustered at the site-level.

[1] Reference group.

*** p<0.001,

** p< 0.01,

* p<0.05,

+ p<0.10.

Table 5

Adjusted means for outcomes adjusted for selection, attrition, and preprogram characteristics, girls (N = 768).

Outcomes	P-3	P-2	P-1	P-K[1]	K-3	None or K only	ES range relative to P-K	F-test
	n = 94	*n* = 199	*n* = 67	*n* = 152	*n* = 70	*n* = 186		
Academic functioning								
Math achievement, grade 3	109.51***	103.16*	101.84	98.51	103.48**	98.97	0.06–0.86	10.15***
Math achievement, grade 8	155.17	152.44	150.63	148.77	150.29	146.36	−0.03–0.55	2.49+
Reading achievement, grade 3	108.90***	100.61*	98.82	95.31	100.06	95.48	0.03–0.83	8.82***
Reading achievement, grade 8	154.19	152.64	150.55	146.43	147.76	146.52	0.02–0.52	2.34+
Social-emotional functioning								
Classroom adjustment, grade 3	21.47	20.07	21.13	19.67	20.18	19.47	−0.04–0.37	2.19+
Classroom adjustment, grade 4–6	21.83	20.08	19.74	19.65	18.89	19.46	−0.16–0.24	1.40
Perceived competence, grades 3–4	29.65	28.63	28.76	28.96	28.69	28.38	−0.16–0.20	0.79
Perceived competence, grades 5–6	35.75	34.74	33.91	34.17	34.26	34.49	−0.06–0.33	0.33
School experiences and expectations								
Ever retained, grades 1–8, %	0.02***	0.14	0.22	0.24	0.14	0.29	−0.59–0.11	30.71***
Years in special education, grade 1–8	0.84	1.15	1.13	1.13	0.42**	1.60	−0.70–−0.25	2.16+
Ever school mobility, grades 4–8, %	0.40**	0.55	0.57	0.62	0.58	0.76*	−0.46–0.29	60.28***
Expectations of college attendance, grade 10, %	0.99**	0.87	0.87	0.88	0.89	0.90	−0.03–0.35	19.43**

Note: Adjusted means/rates on the outcomes are calculated using the marginal means, which were adjusted for preprogram characteristics (i.e., child's gender, child's race, child's age, family variables, and socio-demographic factors), selection and attrition by IPW, and Benjamin-Hochberg correction for multiple comparisons. Standard errors are clustered at the site level.

[1] Reference group.

*** p < 0.001,

** p < 0.01,

* p < 0.05,

+ p < 0.10