Published in final edited form as:
Dev Psychol. 2014 January ; 50(1): 202–215. doi:10.1037/a0032280.

Head Start Participation and School Readiness: Evidence from the Early Childhood Longitudinal Study-Birth Cohort December 10, 2012

RaeHyuck Lee [Doctoral Candidate],
Columbia University School of Social Work 1255 Amsterdam Avenue, New York, NY 10027 rl2507@columbia.edu

Fuhua Zhai [Assistant Professor],
School of Social Welfare, Stony Brook University L2-093 Health Sciences Center Stony Brook, NY 11794 fuhua.zhai@stonybrook.edu

Jeanne Brooks-Gunn [Virginia and Leonard Marx Professor of Child Development and Education],
Teachers College and College of Physicians and Surgeons, Columbia University 525 West 120th Street, New York, NY 10027 brooks-gunn@columbia.edu

Wen-Jui Han [Professor], and
Silver School of Social Work, New York University 1 Washington Square North, New York, NY 10003 wjh3@nyu.edu

Jane Waldfogel [Compton Foundation Centennial Professor for the Prevention of Children and Youth Problems]
Columbia University School of Social Work 1255 Amsterdam Avenue, New York, NY 10027

Abstract

Using data from the Early Childhood Longitudinal Study-Birth Cohort ($n \approx 6{,}950$), a nationally representative sample of children born in 2001, we examined school readiness (academic skills and socio-emotional wellbeing) at kindergarten entry for children who attended Head Start compared to those who experienced other types of child care (prekindergarten, other center-based care, other non-parental care, or parental care). Using propensity score matching methods and OLS regressions with rich controls, we found that Head Start participants had higher early reading and math scores than children in other non-parental care or parental care, but also higher levels of conduct problems than those in parental care. Head Start participants had lower early reading scores compared to children in prekindergarten, and had no differences in any outcomes compared to children in other center-based care. Head Start benefits were more pronounced for children who had low initial cognitive ability or low-educated parents, or attended Head Start for more than 20 hours per week.

Corresponding author: Jane Waldfogel, Columbia University School of Social Work, 1255 Amsterdam Avenue, New York, NY 10027, jw205@columbia.edu.

Keywords

school readiness; ECLS-B; Head Start; Prekindergarten; propensity score matching

Established in 1965 as part of President Lyndon B. Johnson's War on Poverty, Head Start (hereafter, HS) provides education, health, nutrition, and other social services to low-income preschool-age children and their parents. A primary goal of HS is to promote school readiness and recently, the educational aspects of HS's mission have been emphasized through, for example, Congressional mandates to raise the quality of HS teachers (Gormley, Phillips, Adelstein, & Shaw, 2010).

A perennial challenge in HS research is the fact that children served by HS are not a random sample of the population. Several recent observational studies that apply rigorous methods have documented both short- and long-term benefits of HS. In 1998, Congress authorized a random assignment experiment, the Head Start Impact Study (HSIS). This study showed that HS yielded short-term benefits, which, however, diminish by the end of first grade (U.S. Department of Health and Human Services [USDHHS], 2005, 2010).

However, most HS research, including the HSIS, has overlooked another important challenge that could also contribute to divergent findings across studies. Most previous studies have not clearly defined the reference group, although children who do not attend HS may have different types of care arrangements (Lee, Brooks-Gunn, & Schnur, 1988; USDHHS, 2005, 2010; Zhai, Brooks-Gunn, & Waldfogel, 2011). If, as theory would suggest, the effects of HS vary depending on the arrangements to which it is compared, then not taking this into account could obscure those effects and could help explain why studies from different periods or areas produce inconsistent results. The only recent study to directly address this issue showed that the effects of HS did vary by the comparison group (Zhai et al., 2011).

This study advances the literature on the effects of HS on children's school readiness at kindergarten entry in four ways, using data from the Early Childhood Longitudinal Study-Birth Cohort (ECLS-B). First, we clearly define the reference group, comparing HS participants with non-participants who received state-funded prekindergarten (hereafter, pre-K), other center-based care, other non-parental care, or parental care. Second, we examine several aspects of school readiness, including academic skills as well as social and emotional wellbeing. Third, we employ multiple imputation and propensity score matching to address the issues of missing data and selection bias. Finally, we examine whether the associations between HS and children's school readiness differ by child, parent, and program factors (i.e., children's initial cognitive ability, parental education, and the number of hours in HS) to see if benefits are more salient among certain groups of children.

THEORETICAL AND EMPIRICAL BACKGROUND

HS is the single largest publicly funded early childhood education program in the U.S., serving about 900,000 young children from low-income families each year (Office of Head Start, 2010). Most 3- and 4-year-old HS participants are in center-based programs, while the

Early Head Start (EHS) program at younger ages (begun in 1995) has a mix of home- and center-based programs. Eligibility rules restrict admission to children from families with incomes below the poverty line, or to children with disabilities (who can make up not more than 10% of enrolled children). HS is not an entitlement and the availability of slots is limited by the amount of funding. In recent years, HS has served approximately half of eligible children (Currie & Neidell, 2007). HS has always embraced a broad definition of school readiness and thus allocates resources to health, nutrition, and other social services as well as early education.

Theoretical Framework

This study aims to examine whether HS has beneficial links with children's school readiness when compared to other specific types of child care. Our approach is based on Bronfenbrenner's (1979) ecological model, which considers development as a process that unfolds over time and is affected by interactions within different environmental contexts. Our study is also informed by life course developmental psychology, which rests on the premise that trajectories for children at risk can be altered by circumstances and events promoting change over the life course (Brooks-Gunn, 2004; Elder & Shanahan, 2006; Rutter, 2000). Consistent with these perspectives, high quality interactions and learning opportunities within proximal environments such as early childhood educational programs may play a protective role for low-income children by promoting more positive social and academic trajectories and thus increasing the odds of favorable developmental outcomes (Crosnoe, Leventhal, Wirth, Pierce, & Pianta, 2010; Shonkoff & Phillips, 2000).

Thus, we expect that attending a preschool program such as HS, which offers cognitively stimulating interactions with sensitive caregivers, may facilitate cognitive, social, and emotional skills. However, the benefits may vary depending on the setting to which HS is compared. In addition, because environmental experiences interact with children's characteristics (Bronfenbrenner & Ceci, 1994; Bronfenbrenner & Morris, 1998), the effects of preschool programs may be more pronounced for children at higher risk, such as those with low initial cognitive ability or less-educated parents. Finally, to the extent that quantity of experience (i.e., dosage) matters, we would expect longer hours in HS to have larger effects than shorter hours. Prior Research on Head Start and Child Development

A large body of research has examined links between HS and child development. We focus our review on studies of short-term outcomes (i.e., at or before school entry) since that is the topic of our study. Notably, although early studies raised doubts about HS's effectiveness, rigorous observational studies have demonstrated short-term academic benefits (as well as longer-term education and earnings benefits) (e.g., Currie & Neidell, 2007; Currie & Thomas, 1995, 1999; Deming, 2009; Gormley, Phillips, & Gayer, 2008). Comparing children randomly assigned to HS at age 3 or 4 to those who were not, the experimental HSIS found that HS yielded modest but significant short-term impacts on pre-reading, pre-writing, and vocabulary test scores for 3-year-olds, and on pre-reading and pre-writing for 4-year-olds (USDHHS, 2005).

A smaller body of literature has examined the effects of HS on socio-emotional outcomes. Observational studies using rigorous methods to control for selection have yielded mixed

results (Currie & Neidell, 2007; Loeb, Bridges, Bassok, Fuller, & Rumberger, 2007; Magnuson, Ruhm, & Waldfogel, 2007; Zhai et al., 2011). The HSIS found that one-year participation in HS led to a reduction in hyperactive and overall problem behaviors as well as an improvement in parent-reported social competence (USDHHS, 2005).

The reference group in almost all prior studies consisted of all children who did not attend HS, even though care arrangements among non-participants were quite variable. For example, the 4-year-olds in the HSIS control group were in the following settings—another HS center (14%), other centers (35%), other non-parental care (11%), and parental care (40%) (USDHHS, 2005). When HS was introduced in the 1960s, few low-income children attended other types of center-based care. Therefore, the theoretical foundations of HS were based on the assumption that children would benefit from its center-based, educationally-oriented programs, as compared to the inputs they would have received from their parents or other informal arrangements. Today, however, there are more preschool options, and the comparison group now includes more children in other forms of preschool.

A recent study examined whether the associations between HS and school readiness vary according to comparison group (Zhai et al., 2011). Using data on a birth cohort of children in low-income families from 18 large U.S. cities from the Fragile Families and Child Wellbeing Study (FFCWS) and propensity score matching models to account for selection into HS, this study reported that HS participation at age 4 yielded benefits for academic outcomes at age 5 compared to parental care or other non-parental care but not compared to pre-K or other center-based care; HS was also associated with better parent-reported social competence (compared to pre-K, other center-based care, or parental care) and fewer parent-reported behavior problems (compared to other center-based care or other non-parental care).

Research Question and Hypotheses

We compare HS participants with those in other specific types of arrangements, using propensity score matching to reduce selection bias and analyzing a nationally representative birth cohort followed through kindergarten entry. Our primary research question is: do the associations between HS participation and children's school readiness differ depending on the type of child care to which it is compared?

We examine four specific types of care: pre-K, other types of center-based care, other non-parental care, and parental care. Pre-K serves as many children as does HS today. Compared to HS, pre-K usually has a more academic focus (Bellm, Burton, Whitebook, Broatch, & Young, 2002). Two small-scale studies have reported greater academic benefits of pre-K than HS (Gormley et al., 2010; Henry, Gordon, & Rickman, 2006), whereas Zhai et al. (2011) found no differences in academic outcomes. It is not clear whether, or how, we might expect the effects of HS to differ from those of other types of center-based care. HS programs have both advantages and disadvantages with respect to other centers (e.g., they offer more holistic programming including attention to health and nutritional needs, but also tend to have staff with lower educational qualifications). In addition, other center-based programs vary widely in quality (Rigby, Ryan, & Brooks-Gunn, 2007). Prior research found no differences in academic outcomes between HS and other center-based care, but some

advantages of HS in terms of social skills and behavior (Zhai et al., 2011). Compared to informal care or care with parents, HS, as a form of center-based care, should be associated with improved academic competence but also possibly higher levels of behavior problems (Love et al., 2003; National Institute of Child Health and Human Development [NICHD] Early Child Care Research Network [ECCRN], 2003, 2005; Love et al., 2003; Waldfogel, 2006).

In summary, the estimated effects of HS could vary considerably, depending on the comparison groups. Therefore, we have the following two specific hypotheses.

Hypothesis 1: Children who attended HS would have lower levels of academic skills than those in pre-K, similar or higher levels than those in other center-based care, and higher levels than those in other non-parental or parental care.

Hypothesis 2: Children who attended HS would have similar levels of socio-emotional wellbeing than those in pre-K, higher levels than those in other center-based care, and lower levels than those in other non-parental care or parental care. Less is known about the association between type of care and socio-emotional wellbeing so this hypothesis is more exploratory. Differing Associations between Head Start and School Readiness by Subgroups

We further examine whether the associations between HS and school readiness vary by child's initial cognitive ability, parental education, or the number of hours per week in HS. We expect children who have lower initial cognitive ability or whose mothers have low levels of education will benefit more from HS's educational programming. We also expect children who attend HS for more hours will benefit more than those who attend for fewer hours. This might be particularly relevant for the comparison with the more academically focused pre-K.

Prior research with Head Start has found greater gains from HS for participants with lower cognitive skills (Lee et al., 1988). The HSIS also found that benefits of HS in socio-emotional wellbeing were more salient among children with low academic skills at baseline (USDHHS, 2010). With respect to differing effects of early childhood intervention by parental education, studies of other interventions have found greater effects for children of less educated parents (e.g., Brooks-Gunn, Gross, Kraemer, Spiker, & Shapiro, 1992; the Infant Health and Development Program [IHDP], 1990). Regarding the association between the amount of exposure and children's school readiness, benefits of greater exposure have been reported for IHDP (Hill, Brooks-Gunn, & Waldfogel, 2003) as well as HS (Wen, Leow, Hahs-Vaughn, Korfmacher, & Marcus, in press). Magnuson et al. (2004) found greater gains on academic skills for children who attended HS or pre-K for more than 20 hours per week than those who attended those programs for 20 hours or less per week, while Loeb et al. (2007) reported that children who attended center care for 30 hours or more per week showed better reading skills but also more behavior problems than those who attended center care for fewer hours. Based on this empirical evidence, we add a third hypothesis:

Hypothesis 3: HS participation would be more beneficial for school readiness for children who had lower initial cognitive ability and less-educated parents than for their counterparts, while full-time HS (more than 20 hours per week) will be associated with higher levels of

academic skills but also more behavior problems than part-time HS (20 hours or less per week). The Present Study

Our study adds to the evidence provided in Zhai et al. (2011) but extends it in the following ways. First, we use a nationally representative birth cohort study, which assesses child development at ages 9, 24, and 48 months. Second, we use a sample of children from rural, urban, and suburban areas, whereas the Zhai et al. study's data included cities only. Third, we examine both math and reading scores (not just language). Fourth, we use socio-emotional outcomes reported by teachers, which may be more relevant for school functioning than those reported by parents (Konold & Pianta, 2007; Waterman, McDermott, Fantuzzo, & Gadsden, 2012). Fifth, we conduct moderating analyses by children's early cognitive skills, parental education, and the number of hours in HS.

METHOD

Data

The data for this study come from the ECLS-B, a nationally representative sample of approximately 10,700 children born in the U.S. in 2001 and followed longitudinally from birth to kindergarten (Nord et al., 2004).[1] The first interview (including a parent survey and direct child assessment) was conducted at 9 months after birth, followed by interviews when the child was approximately 2, 4 (preschool wave), and 5 or 6 (kindergarten wave) years old.[2] The data used in this study come from all four waves. HS participation information was collected from the preschool survey, with child outcomes from the kindergarten survey, and other covariates including pre-treatment scores from the 9-month and/or 2-year surveys.

The original sample was reduced to 7,700 cases at the kindergarten wave due to financial constraints (in all of our analyses, we used weights ['WK1R0'] provided by the ECLS-B to adjust for this sample reduction) (Snow et al., 2009). Of the 7,700, about 7,000 parents participated in the parent interview and about 6,900 children took part in the assessment. Our kindergarten analysis sample was further reduced by excluding about 50 children who had missing information on child care arrangements at preschool or outcome measures at kindergarten (under 1% of the original sample and 1% of the preschool sample), yielding about 6,950 cases for our analysis (65% of the original sample and 90% of the kindergarten sample) (as detailed below, we used multiple imputation to address missing data for covariates).

We compared children in our analytic sample with the children excluded due to attrition or missing data on child care arrangements or outcome measures. Children in our sample were slightly more advantaged than those excluded: for example, more likely to have highly educated parents and high-income families. However, these differences tended to be modest, and our analytic sample was similar to the original ECLS-B sample on most demographic variables. Outcome Measures

[1]In accordance with U.S. Department of Education requirements, all reported samples sizes are rounded to the nearest 50.
[2]Aproximately 75% of the children entered kindergarten in fall 2006 and had their kindergarten survey that year. The remaining children who entered kindergarten in 2007 were assessed in 2007. Thus, we obtained information about children's kindergarten outcomes by combining the 2006 and 2007 surveys.

We analyzed six outcome measures, all assessed at kindergarten entry, including three measures of academic school readiness and three of socio-emotional wellbeing.[3] To facilitate comparison of effect sizes across outcome variables, we standardized all outcomes (in the original sample) to have a mean of 0 and a standard deviation of 1.

Expressive language—The Let's Tell Stories subset of Preschool Language Assessment Scales (PreLAS) was utilized to measure children's language skills (Duncan & De Avila, 1998). The scale consists of reading two stories to a child, recording the child's response, and rating the response by trained field coders with a range from 0 ("no response") to 5 ("articulate, detailed sentences, vivid vocabulary, and complex constructions") (Snow et al., 2009). The average reliabilities (percent of agreement) across all field coders for both stories were 72% for the 2006 kindergarten survey and 95% for the 2007 kindergarten survey (Najarian, Snow, Lennon, & Kinsey, 2010). The present study used the average score of both stories, provided in the ECLS-B data set. Higher scores reflect higher levels of expressive language skills.

Early reading—Children's reading ability was assessed using 60 receptive and literacy items developed by the ECLS-B (7 for English language skills/oral language, 15 for phonological awareness, 14 for letter and letter-sound knowledge, 6 for print conventions, 11 for word recognition, and 7 for vocabulary) (Najarian et al., 2010; Snow et al., 2009). The internal consistency across these items in the original sample was .92 for the 2006 kindergarten wave and .93 for the 2007 kindergarten wave (Najarian et al., 2010).

Mathematics—Children's math ability was measured with 58 items developed by the ECLS-B (41 for number sense, properties, and operations; 3 for measurement, 4 for geometry and spatial sense; 3 for data analysis, statistics, and probability; and 7 for patterns, algebra, and functions) (Najarian et al., 2010; Snow et al., 2009). The internal consistency across all items in the original sample was .92 for both 2006 and 2007 kindergarten waves (Najarian et al., 2010).

Conduct problems—To measure conduct problems, following Bradbury, Corak, Waldfogel, and Washbrook (2011), we used four items that asked the teacher about children's temper tantrums, aggressiveness, annoying, and destructive behavior with a 5-point Likert scale (1 = "never" to 5 = "very often"). We created a standardized total score by summing and standardizing all four items (α = .87). Higher scores indicate higher levels of conduct problems.

Hyperactivity/Inattention—Following Bradbury et al. (2011), hyperactivity/inattention problems were assessed using the total score of 6 teacher-reported items (α = .92) with a 5-point Likert scale (1 = "never" to 5 = "very often"): acting impulsively, being overly active, being fidgety, having difficulty concentrating, paying attention well (reverse coded), and

[3]We only present results for early reading, mathematics, and conduct problems in the results tables, because we found very little evidence of effects on the other three measures—expressive language, hyperactivity/inattention, and pro-social behaviors (available upon request).

keeping working until finished (reverse coded). Higher scores mean more problematic behavior.

Pro-social behavior—Ten items reported by the teacher with a 5-point Likert scale (1 = "never" to 5 = "very often") were used to measure children's pro-social behavior. The items include statements such as "the child is invited by other children to play," "the child shows eagerness to learn new things," "the child is accepted and liked by other children," and "the child comforts other children who are upset." We constructed a composite score by summing all items (α = .87). Higher scores indicate more pro-social behaviors.

Correlations among outcomes—Strong correlations were found between reading and math scores (r = 0.80), between conduct problems and hyperactivity/inattention (r = 0.74), and between hyperactivity/inattention and pro-social behavior (r = -0.57). The rest of the correlation coefficients showed weak or modest relationships (ranging from r = -0.46 to r = 0.37).

Child Care Arrangements at Preschool

We obtained information on child care arrangements from the preschool survey, administered when children were approximately 4-years-old. Parents were asked a set of questions about types of child care including HS and other center-based care (day care centers, nursery schools, preschool programs, or pre-K programs). Based on these questions, we created mutually exclusive categories following prior research (Magnuson et al., 2007; Zhai et al., 2011). Parents were first asked if their children were currently attending HS on a regular basis. We considered children whose parents answered 'yes' as HS participants. Second, for parents answering 'no,' they were then asked if their children were presently attending a day care center, nursery school, preschool, or pre-K program on a regular basis (not including HS programs). For children whose parents answered 'yes' to this question, based on the information about the center-based program where the child spent the most time, we divided them into two groups: pre-K, or other center-based care (day care center, nursery school, or preschool program). Next, if the answer about participation in center-based care was 'no,' parents were further asked if their child was currently receiving care from a relative other than a parent, or a non-relative, on a regular basis. If parents answered 'yes' and their children received other non-parental care for at least 8 hours per week for a month or more, we categorized those children as receiving other non-parental care. Finally, children were categorized as receiving parental care if they were neither regularly attending child care centers (i.e., HS, pre-K, or other center-based care) nor receiving other non-parental care for at least 8 hours per week for a month or more.[4]

Using the above categorization, at the preschool wave, 15.8% of children were attending HS, 15.2% in pre-K, 37.7% in other center-based care, 10.1% in other non-parental care, and

[4]Just under half (42%) of the HS participants also spent some time in other care arrangements. Because our focus is on HS, we coded children as HS participants if they participated in HS on a regular basis, regardless of whether this was the arrangement where they spent the most time. In supplemental analyses, we examined whether the effects of HS varied depending on whether children attended only HS or HS combined with some other arrangements. Briefly, in results not shown but available on request, we found that children in the former group showed almost the same results as in the main models, whereas children in the latter group showed only a few differences.

21.2% in parental care. To compare HS with the other specific types of child care arrangements, we conducted four sets of analysis: HS versus pre-K, HS versus other center-based care, HS versus other non-parental care, and HS versus parental care. Although our primary focus was on HS versus other specific care arrangements at the preschool wave, we also created variables to control for earlier child care arrangements. Specifically, we created a set of dummy variables indicating parental care, relative care, non-relative care, and center-based care for children's care arrangements at 9 months and 2 years, and included these variables in our analyses.

Other Covariates

Based on the ecological theory (Bronfenbrenner, 1979) as well as previous studies (e.g., Currie & Thomas, 1995; Zhai et al., 2011), we controlled for a set of covariates measured at the 9-month and/or 2-year assessments, which might be associated both with children's HS participation and their school readiness. Child characteristics (at the 9-month survey unless otherwise noted) included gender, age in months, race/ethnicity, low birth weight (<2.5kg), prematurity (gestation weeks <37), multiple births, duration of breast-feeding, number of siblings, and health status (reported by the mother at the 2-year wave). Parental and family characteristics (based on information collected at the 9-month survey) included maternal age at child birth, whether the mother lived with her biological mother and father until 16 years of age, mother's marital status at birth (married vs. unmarried), mother's place of birth (foreign- vs. native-born), primary home language (English vs. non-English), parental education (whichever parent had the highest education degree), parental occupational prestige (whichever parent had the highest prestige score), and maternal depressive symptoms (constructed using the Center for Epidemiological Studies–Depression Scale [CES-D; Radloff, 1977], and averaging the total scores [α = .88] of 12 items that ranged 0 to 36 with higher scores indicating higher level of depression symptoms). Maternal and family covariates collected at the 2-year wave included maternal employment status (full-time, part-time, or not working), mother's self-rated health status (excellent, very good, good, or fair/poor), urbanicity (urban vs. rural), region (Northeast, Midwest, South, or West), receipt of Food Stamps, receipt of WIC, and receipt of TANF.

Five parenting behavior variables were also included as covariates. First, the Knowledge of Infant Development Inventory (KIDI; MacPhee, 1981) collected at the 9-month survey measures parental beliefs about norms and behaviors that can affect infant growth. We summed up the total scores of 11 items that ranged from 0 to 11 (α = .59), with higher scores indicating greater knowledge of infant-rearing behaviors. Second, following Rodriguez et al. (2009), mother's cognitively stimulating activities at the 2-year wave were constructed by taking the total score of the three items: reading books, telling stories, and singing songs, with a total score ranging from 3 to 12 (α = .59). Third, mother's spanking behavior at the 2-year wave was measured as a binary variable with a value of 1 if she never spanked her children and 0 otherwise. Two family routine variables at the 2-year wave were measured using the questions asking the mother whether she had a routine for regular bedtime and how often her family ate dinner together per week. Also, two indicators for the presence of the father at the 9-month and 2-year waves were included.

Pre-Treatment Cognitive and Socio-Emotional Scores

Most children begin to attend HS at age 3 or 4. Thus, in supplemental models, we used outcomes data from the 2-year survey to control for possible pre-existing differences in children's pre-treatment scores (i.e., their level of development prior to entry into HS) (Lee et al., 1998; McCartney, Burchinal, & Bub, 2006). Specifically, we used the total mental scale score (a reliability of 0.88; Andreassen & Fletcher, 2007) of the Bayley Short Form-Research Edition (BSF-R; Bayley, 1993; Nord et al., 2004) as a pre-treatment cognitive outcome and the total score (7 individual items; $\alpha = .66$) of the Infant/Toddler Symptoms Checklist (ITSC; Nord et al., 2004) as a pre-treatment socio-emotional outcome.

Missing Information on Covariates

For all analyses, we used multiple imputation to address missing data on covariates (Rubin, 1987). Most covariates had low missing rates (ranging from 0% to 8%), although some covariates (e.g. scales such as BSF-R) had slightly higher rates (ranging from 8% to 10%). We used the ICE command in Stata to create five imputed data sets (Royston, 2005). We then used MICOMBINE to combine the imputed five data sets, estimate separate regressions for each one, and then obtain final parameter estimates reflecting averages across the regressions.

Analytic Approaches

To address potential selection bias, we specified four different models (two OLS models with rich controls and propensity score matching without and with pre-treatment scores). In addition, all four models were conducted for four sets of analyses (i.e., HS vs. pre-K, HS vs. other center-based care, HS vs. other non-parental care, and HS vs. parental care) to compare HS participants to specific groups of children. We begin with two OLS regressions with rich controls. Specifically, Model 1 controls only for HS participation and child demographic covariates, while Model 2 adds parental and family characteristics. However, OLS models, even with the inclusion of rich controls, may still yield biased estimates because of omitted variables. To address this issue, this study used a propensity score matching approach. Considering selection into the treatment as a function of all possible observed covariates, this method makes treatment and control groups as similar as possible on the observables and then compares the outcomes of both groups (Dehejia & Wahba, 2002; Rosenbaum & Rubin, 1983, 1985). Including theoretically relevant covariates can further improve bias reduction (Heckman, Ichimura, & Todd, 1997, 1998). However, it is important to note that propensity score matching methods can not adjust for unobserved differences between treatment and controls (Hill et al., 2005).

This study carried out propensity score matching in three steps (see the supplemental online material for additional information). First, a probit regression was conducted to predict propensity scores for each observation, that is, the probability of being in HS compared to each other type of care, including all covariates controlled for in Model 2 and weighted using the sampling weight ('WK1R0'). Second, for each comparison of HS with other care arrangements, we employed the PSMATCH2 command in Stata with the radius matching option to match each HS participant with all of the comparison units within a caliper (Dehejia & Wahba, 2002). The common support option was used to limit observations to

those whose propensity scores fall within the common support region (Gibson-Davis & Foster, 2006; Leuven & Sianesi, 2003). Before conducting the matching procedure, we randomly sorted data to prevent cases being affected by the order of the observations. Propensity score matching in each imputed data set yielded well-balanced treatment and control groups in comparisons of HS with each of the other specific care groups, eliminating all statistically significant differences in covariates between groups (see Table A in the supplemental online material). In addition, although the distributions of propensity scores between the matched treatment and comparison groups were not perfectly symmetric, both groups showed sufficient common support, which means the matched treatment group in each comparison had a reliable counterpart.

In the third step, regression-adjusted differences were employed to improve the precision of estimates. Specifically, Model 3, with the same covariates as Model 2, was estimated using the weights generated from the matching process as probability weights in OLS regressions. These estimates improve upon the prior OLS estimates in terms of addressing potential selection bias by placing more weight on control cases that are most similar to the HS participants, and less weight on those that are least similar. Model 4 was identical to Model 3 except that it also included controls for the child's scores on relevant outcomes prior to the child's enrollment in HS. The inclusion of such pre-treatment scores can control for any further heterogeneity not already accounted for with the rich controls and propensity score matching.

Finally, we conducted sub-group analyses by children's initial cognitive ability (low ability was defined as a standardized BSF-R score at 9 months at -0.50 standard deviations or less vs. higher) and parental education (children of parents with a high school diploma or less vs. children of parents with more than a high school diploma). Also, we explored if the associations between HS and school readiness differ by whether the child attended HS full-time (more than 20 hours per week) or part-time (20 hours or less per week).

RESULTS

Descriptive Statistics and Results of Propensity Score Matching

We first examined children's school readiness outcomes and pre-treatment covariates by specific types of care arrangements (Table 1). Prior to propensity score matching, HS participants tended to have lower scores in all three cognitive outcomes and pro-social behaviors, and higher scores in conduct problems and hyperactivity/inattention, compared to children in other types of care, except that HS participants showed similar scores in pro-social behaviors to children who received other non-parental care. Gaps were larger between HS participants and children who attended pre-K or other center-based care, and smaller between HS participants and children in other non-parental or parental care. These differences by child care group show the importance of clearly defining the comparison group. Table 1 also reveals many significant differences in characteristics between HS and non-HS participants, indicating substantial selection into HS versus other forms of care. After propensity score matching, no significant mean differences were detected between HS participants and non-participants.

Associations between Head Start and School Readiness

Using four model specifications as detailed above, the associations of HS with school readiness outcomes compared to four other specific care arrangements were estimated (Table 2). Because all outcome variables were standardized with a mean of 0 and a standard deviation of 1, the estimated coefficients are interpreted as effect sizes (proportion of a standard deviation [SD]). In discussing results, we focus on statistically significant effects (at $p < .05$ level).

Many significant differences between HS versus other arrangements are seen in Model 1. Those differences are reduced by the inclusion of maternal and family characteristics (Model 2) and by propensity score matching (Model 3). After propensity score matching (Model 3), HS participants tended to have lower early reading scores than comparable children in pre-K (-0.18 *SD*s) but tended to have similar scores in all outcome measures to children who attended other center-based care (in Panels A and B, respectively). In contrast, HS participants tended to have higher early reading and math scores compared to children who had other non-parental care (0.12 *SDs* and 0.11 *SDs*, respectively, in Panel C) and compared to those in parental care (0.08 *SDs* and 0.09 *SDs*, respectively, in Panel D). Also, HS participants had higher scores in conduct problems (0.12 *SDs*) than those of parental care (Panel D). Model 4 in Table 2 is the same as Model 3 but adds controls for pre-treatment scores. The results show that including these scores did not make notable changes in the estimated effects of HS participation.

Do the Associations Differ by Subgroup?

Table 3 shows the results of subgroup analyses, based on Model 3. As expected, benefits of HS participation were more positive for children with low initial cognitive ability than for their counterparts, but this was only found when compared to pre-K. Compared to pre-K, HS was associated with significantly lower scores in reading and math for children with high initial cognitive ability, but with higher scores in math (and similar scores in early reading) for children with low initial cognitive ability (Panel A of Table 3).

As hypothesized, HS participation was more beneficial to children's school readiness for children with less-educated parents than for those with more-educated parents. Specifically, compared to children in pre-K or other center-based care, HS participants had significantly lower scores in early reading for children with more-educated parents but showed no differences for children with less-educated parents (Panels A and B of Table 3). Similarly, compared to parental care, HS was associated with significantly higher reading and math scores for children with less-educated parents, whereas there were no differences in those outcomes for children with more-educated parents (Panel D of Table 3). Also, compared to children in other center-based care, HS participants showed more conduct problems for those with more-educated parents but not for those with less-educated parents (Panel B of Table 3).

Finally, as hypothesized, full-time HS was associated with better academic skills. Compared to children in pre-K, part-time HS participants had significantly lower scores in early reading, whereas full-time HS participants showed no difference (Panel A of Table 3). Also,

full-time HS participants showed significantly higher scores in early reading (compared to children in other-non parental care or parental care) and math (compared to children in other non-parental care) while part-time HS participants did not (Panels B and C of Table 3). In addition, as expected, full-time HS was associated with higher levels of behavioral problems. Full-time HS participants had more conduct problems compared to children in pre-K or parental care, whereas part-time HS participants showed no differences (Panels A and D of Table 3).

DISCUSSION AND CONCLUSION

Using a large contemporary and nationally representative sample and a well-constructed non-experimental research design, this study tested three hypotheses on how HS children compared to children who attended other types of child care in terms of several dimensions of school readiness (the first two hypotheses) and whether the estimated differences varied by children's initial cognitive ability, parental education, and the number of hours in HS (the third hypothesis). Building upon the theoretical frameworks and prior literature, this study tried to detect more precise associations between HS participation and school readiness by clearly defining the comparison group, rather than comparing HS participants to all other children.

As hypothesized, since children interact with different environments in different programs, substantially differing associations of HS with school readiness outcomes were found depending on the type of care to which HS was compared. The main findings indicate that HS participants had higher early reading and math scores than children who received other non-parental care or parental care (but more behavior problems than children in parental care), lower early reading scores than children in pre-K, and similar scores to children in other types of center care. Also as hypothesized, since children's developmental trajectories are affected by interactions with environments as well as individual characteristics, children with low initial cognitive ability, with less-educated parents, and who attended HS full-time tended to benefit academically more from HS than their counterparts; however, children with more-educated parents and who attended HS full-time tended to have higher levels of behavior problems.

We started this study with the working hypothesis that early childhood interventions can alter developmental trajectories of poor children in a positive direction, and found empirical support for the conclusion that HS participants did have better cognitive development compared to non-participants, particularly those in other non-parental care or parental care. This result suggests that an early childhood intervention can play a role in shaping early childhood development, but that role may depend on the counterfactual. If researchers and policymakers want to answer the question of whether HS works, it is essential to extend that question to ask "compared to what?"

HS showed no differences with other center-based care (which included day care centers, preschools, and nursery schools), but as we hypothesized, children who had attended pre-K had better cognitive development in kindergarten compared to HS participants. Children in pre-K may have more academically-focused environments than those in HS, reaffirming our

assumption that benefits of early intervention depend on what occurs within that setting. In addition, the significant differences between these two programs may derive from pre-K programs' higher quality in structural indicators, such as levels of teacher education (Bellm et al., 2002). Over half (57%) of teachers in pre-K programs have a college degree (Barnett, Carolan, Fitzgerald, & Squires, 2011), in contrast to only 30% of HS teachers (30% of whom have Child Development Associate or alternative credentials) (Hart, 2005). Of note, the positive effects of pre-K relative to HS were found in early reading, but not math. While it is not clear why this would be the case, it could be that pre-K programs were more literacy focused during the period of time when the ECLS-B data were collected, considering that there were several literacy/early reading initiatives during this period of time (e.g., Early Reading First) and several District and Statewide assessment initiatives focused on early literacy screening at the beginning of kindergarten (e.g., Dynamic Indicators of Early Literacy Skills).

Consistent with theoretical expectations, this study showed that HS yielded academic benefits compared to other non-parental care or parental care. The benefits were modest (roughly 0.08 to 0.12 of a standard deviation) but consistently found for both early reading and math.

Whereas the Zhai et al. (2011) study using parent-reported behavior found enhanced social competence and reduced attention problems among HS participants compared to all non-participants, our study using outcomes reported by teachers found more conduct problems among the participants. In addition, our study found that HS participants had more teacher-reported conduct problems than children in parental care whereas the Zhai et al. study found no differences in parent-reported externalizing problems between the two groups. Our result is consistent with prior research finding links between center-based care and child behavior problems reported by teachers at school entry (e.g., Belsky et al., 2007) but is of concern given HS's focus on social aspects of school readiness and given some evidence that high-quality child care programs are associated with lower levels of behavior problems (e.g., Brooks-Gunn et al., 1994; Love et al., 2003;). This result suggests that quality improvements to the HS program might be needed to improve participants' socio-emotional development.

There are a few other differences between our results and those of the recent study by Zhai et al. (2011). In general, our estimates of the associations between HS and academic school readiness outcomes are similar but tend to be smaller. While we can only speculate, one possible explanation might be the difference in outcome measures. Zhai et al. analyzed data from the FFCWS, which used the original PPVT and WJ scales, while the measures in our study were assessments specifically developed by the ECLS-B and scored using Item Response Theory (IRT) methods. Compared to effect sizes from only one scale (e.g., PPVT), those from combined scores of diverse direct assessment scales might be smaller if there were respondents who score well on some scales but poorly on the others. Another possible explanation might be the difference in samples. Our study used a nationally representative sample from the ECLS-B, whereas Zhai et al. used an exclusively urban sample from the FFCWS. This difference might contribute to the variation in effect sizes if HS is more effective in urban areas. Furthermore, the ECLS-B oversampled twins and low birth weight babies whose early development tends to be delayed, whereas the FFCWS did

not. Finally, it is possible that, if norms for the assessments are based on a national sample, low-income children may show some restriction in range on these assessments compared to the national sample (McDermott et al., 2009).

The HSIS's follow-up analysis found that overall, sustained effects were not detected when children entered kindergarten and first grade (USDHHS, 2010). Our findings might shed some light on those results. We found no significant differences in kindergarten outcomes between HS participants and nonparticipants when we did not take their alternative care arrangements into account (results not shown but available upon request). However, when we disaggregated nonparticipants by type of arrangement, the present study showed that HS participants did have better academic outcomes than children who had used parental care or other non-parental care, but similar or worse than children who had attended pre-K or other center-based care.

The fact that we found the largest academic benefits for children who would otherwise not have attended any center-based care is both theoretically- and policy-relevant, given that many eligible children who do not participate in HS do not get into any center-based care arrangements. The HSIS (USDHHS, 2005) showed that, among children in the control group, 57% of the 3-year-olds and 52% of the 4-year-olds did not attend any type of center-based care. Indeed, 39% of the 3-year-olds and 42% of the 4-year-olds in the control group received only parental care. Our findings suggest that HS participation is particularly effective for these children, who are most at risk for entering kindergarten behind their peers.

The identification of children who might obtain the greatest gains from HS participation is further bolstered by the findings from our subgroup analyses. We found that HS participation tended to be more beneficial than pre-K in developing academic skills for children with low initial cognitive ability, which is notable given that previous studies have found that pre-K programs were more effective in boosting children's cognitive development than HS programs (Gormley et al., 2010; Henry et al., 2006; Magnuson et al., 2007). One possible explanation is that HS's curriculum is more tailored to the neediest children. For example, HS's curriculum includes developmentally appropriate learning practices (e.g., learning basic skills) (Lee et al., 1988; Taylor, 2000), which might not be included in pre-K's curriculum with the assumption that children had already attained those skills.

As hypothesized, we also found that HS participation provided more benefits for children of less-educated parents than for those of more-educated parents. Compared to pre-K or other center-based care, children of less-educated parents obtained at least similar gains from HS participation, and they showed better academic ability in the comparison of HS with parental care. Children from less advantaged backgrounds may be less likely to experience educational home environments (e.g., cognitively stimulating activities) than their counterparts (Bradley, Corwyn, McAdoo, & Garcia Coll, 2001; Linver, Brooks-Gunn, & Kohen, 2002).

In addition, as expected, participating in HS full-time was associated with higher academic skills as well as more behavior problems than participating in HS part-time. Children in full-time HS had comparable academic scores to children in pre-K but higher behavior problem scores, perhaps suggesting that more attention to social and emotional wellbeing, or more generally to higher quality care, might be needed (Love et al., 2003; McCartney et al., 2010).

This study possesses several limitations. First, as is common with survey data, parent-reported care arrangements may not always be accurate, which would bias estimated effects. Second, our estimates may be subject to omitted variable bias, since even the propensity score matching models cannot account for unobservables that may affect both children's participation in HS and their outcomes. Third, we lack information on HS program quality, an important omission given that the effects of HS participation likely vary depending on the quality of the program.

In spite of these limitations, the findings of this study contribute to our understanding of the associations between children's experiences in child care and their early development. Furthermore, the findings may suggest policy implications. First, HS funding, where possible, should be allocated to target children who would receive other non-parental care or parental care if not attending a HS program, because the cognitive benefits to children in these groups are the largest. More generally, expanding both HS opportunities as well as state-funded preschool and pre-K programs to all poor children should be a priority. Second, policy makers need to pay more attention to the quality of HS and other programs, especially as linked to curriculum on literacy, math, and science (Barnett, 2011) as well as cognitively stimulating and supportive interactions between teachers and children (Mashburn & Pianta, 2010). Finally, the evidence that full-time HS programs were more beneficial for children's academic school readiness than part-time programs should be of high interest to policy makers.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We gratefully acknowledge support from award number R01HD047215-05 and R24HD058486 from the Eunice Kennedy Shriver National Institute of Child Health & Human Development (NICHD). The content is solely the responsibility of the authors and does not necessarily represent the official views of NICHD or the National Institutes of Health.

REFERENCES

Andreassen, C.; Fletcher, P. Early Childhood Longitudinal Study, Birth Cohort (ECLS–B) Psychometric Report for the 2-Year Data Collection (NCES 2007–084). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education; Washington, DC: 2007.

Bagner DM, Pettit JW, Lewinsohn PM, Seeley JR. Effect of maternal depression on child behavior: A sensitive period? Journal of the American Academy of Child & Adolescent Psychiatry. 2010; 49(7): 699–707. [PubMed: 20610139]

Barnett W. Effectiveness of early educational intervention. Science. 2011; 333(975):975–978. [PubMed: 21852490]

Barnett, W.; Carolan, ME.; Fitzgerald, J.; Squires, JH. The state of preschool 2011. The National Institute for Early Education Research (NIEER); Brunswick, NY: 2011.

Bayley, N. Bayley Scales of Infant Development, Second edition manual. The Psychological Corporation; San Antonio, TX: 1993.

Bellm, D.; Burton, A.; Whitebook, M.; Broatch, L.; Young, MP. Inside the pre-K classroom: A study of staffing and stability in state-funded prekindergarten programs. Center for Child Care Workforce; Washington, DC: 2002.

Belsky J, Vandell DL, Burchinal M, Clarke-Stewart KA, McCartney K, Owen MT, NICHD Early Child Care Research Network. Are there long-term effects of early child care? Child Development. 2007; 78:681–701. [PubMed: 17381797]

Bradbury, B.; Corak, M.; Waldfogel, J.; Washbrook, E. Inequality during the early years: Child outcomes and readiness to learn in Australia, Canada, United Kingdom, and United States. 2011. IZA Discussion Paper Series No. 6120

Bradley RH, Corwyn RF, McAdoo HP, Garcia Coll C. The home environments of children in the United States Part I: Variations by age, ethnicity, and poverty status. Child Development. 2001; 72:1844–1867. [PubMed: 11768149]

Bronfenbrenner, U. The ecology of human development: Experiments by nature and design. Harvard University Press; Cambridge, MA: 1979.

Bronfenbrenner U, Ceci SJ. Nature-nurture reconceptualized in developmental perspective: A bioecological model. Psychological Review. 1994; 101:568–586. [PubMed: 7984707]

Bronfenbrenner, U.; Morris, PA. The ecology of developmental processes. In: Lerner, R., editor. Handbook of child psychology: Vol. 1. Theoretical models of human development. 5th ed.. Wiley; New York: 1998. p. 993-1028.

Brooks-Gunn, J. Intervention and policy as change agents for young children. In: Chase-Lansdale, PL.; Kiernan, K.; Friedman, RJ., editors. Human development across lives and generations: The potential for change. Cambridge University Press; New York: 2004. p. 293-340.

Brooks-Gunn J, Gross RT, Kraemer HC, Spiker D, Shapiro S. Enhancing the cognitive outcomes of low birth weight, premature infants: For whom is the intervention most effective? Pediatrics. 1992; 98(6):1209–1215. [PubMed: 1375731]

Brooks-Gunn J, Han W-J, Waldfogel J. First-year maternal employment and child development in the first seven years. Monographs of the Society for Research in Child Development (SRCD). 2010; 75(2):1–147.

Brooks-Gunn J, McCarton C, Casey P, McCormick M, Bauer C, Bernbaum J, Meinert CL. Early intervention in low birth weight, premature infants: Results through age 5 years from the Infant Health and Development Program. Journal of the American Medical Association. 1994; 272:1257–1262. [PubMed: 7933370]

Crosnoe R, Leventhal T, Wirth RJ, Pierce KM, Pianta RC. Family socioeconomic status and consistent environmental stimulation in early childhood. Child Development. 2010; 81(3):972–987. [PubMed: 20573117]

Currie J, Neidell M. Getting inside the ‘Black Box’ of Head Start quality: What matters and what doesn’t. Economics of Education Review. 2007; 26(1):83–99.

Currie J, Thomas D. Does Head Start make a difference? American Economic Review. 1995; 85:341–364.

Currie J, Thomas D. Does Head Start help Hispanic children? Journal of Public Economics. 1999; 74:235–262.

Dehejia RH, Wahba S. Propensity score-matching methods for nonexperimental causal studies. Review of Economics and Statistics. 2002; 84(1):151–161.

Deming D. Early childhood intervention and life-cycle skill development: Evidence from Head Start. American Economic Journal: Applied Economics. 2009; 1(3):111–134.

Duncan, SE.; De Avila, EA. PreLAS 2000. CTB/McGraw-Hill; Monterey, CA: 1998.

Elder, GH.; Shanahan, MJ. The Life Course and Human Development. In: Damon, W.; Lerner, RM., editors. Handbook of Child Psychology, Volume 1, Theory. U Chicago Press; 2006. p. 665-715.

Gibson-Davis C, Foster E. A cautionary tale: Using propensity scores to estimate the effect of Food Stamps on food insecurity. Social Service Review. 2006; 80(1):93–126.

Gormley WT, Phillips D, Adelstein S, Shaw C. Head Start's comparative advantage: Myth or reality? Policy Studies Journal Volume. 2010; 38(3):397–418.

Gormley WT, Phillips D, Gayer T. Preschool programs can boost school readiness. Science. 2008; 320:1723–1724. [PubMed: 18583597]

Hart, K. Making the case: Improving Head Start teacher qualifications requires increased investment. The Center for Law and Social Policy; Washington, DC: 2005.

Heckman JJ, Ichimura H, Todd TE. Matching as an econometric evaluation estimator: Evidence from evaluating a job training program. Review of Economic Studies. 1997; 64(4):605–654.

Heckman JJ, Ichimura H, Todd TE. Matching as an econometric evaluation estimator. Review of Economic Studies. 1998; 65(2):261–294.

Henry GT, Gordon CS, Rickman DK. Early education policy alternatives: Comparing quality and outcomes of Head Start and State Prekindergarten. Educational Evaluation and Policy Analysis. 2006; 23(1):77–99.

Hill LJ, Brooks-Gunn J, Waldfogel J. Sustained effects of high participation in an early intervention for low-birth-weight premature infants. Developmental Psychology. 2003; 39:730–744. [PubMed: 12859126]

Hill LJ, Waldfogel J, Brooks-Gunn J, Han W. Maternal employment and child development: A fresh look using newer methods. Developmental Psychology. 2005; 41:833–850. [PubMed: 16351331]

Konold TR, Pianta RC. The influence of informants on ratings of children's behavioral functioning. Journal of Psychoeducational Assessment. 2007; 25(3):222–236.

Lee VE, Brooks-Gunn J, Schnur E. Does Head Start work? A 1-year follow-up comparison of disadvantaged children attending Head Start, no preschool, and other preschool programs. Developmental Psychology. 1988; 24:210–222.

Leuven, E.; Sianesi, B. [Retrieved April 14, 2011] PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing. 2003. from http://ideas.repec.org/c/boc/bocode/s432001.html

Linver MR, Brooks-Gunn J, Kohen D. Family processes as pathways from income to young children's development. Developmental Psychology. 2002; 5:719–734. [PubMed: 12220050]

Loeb S, Bridges M, Bassok D, Fuller B, Rumberger R. How much is too much? The influence of preschool centers on children's social and cognitive development. Economics of Education Review. 2007; 26(1):52–66.

Love JM, Chazan-Cohen R, Raikes H, Brooks-Gunn J. What makes a difference? Early Head Start Evaluation findings in a developmental context. Monographs of the Society for Research in Child Development (SRCD). in press.

Love JM, Harrison L, Sagi-Schwartz A, van IJzendoorn MS, Ross C, Ungerer JA, Chazan-Cohen R. Child care quality matters: How conclusions may vary with context. Child Development. 2003; 74:1021–1033. [PubMed: 12938696]

Love JM, Kisker EE, Ross C, Raikes H, Constantine J, Boller K, Vogel C. The effectiveness of Early Head Start for 3-year-old children and their parents: Lessons for policy and programs. Developmental Psychology. 2005; 41:885–901. [PubMed: 16351335]

MacPhee, D. Manual: Knowledge of Infant Development Inventory. Department of Psychology, Unpublished manuscript; Chapel Hill, NC: 1981.

Magnuson K, Ruhm C, Waldfogel J. Does prekindergarten improve preparation and performance? Economics of Education Review. 2007; 26(1):33–51.

Magnuson KA, Meyers M, Ruhm CJ, Waldfogel J. Inequality in preschool education and school readiness. American Educational Research Journal. 2004; 41:115–157.

Mashburn, AJ.; Pianta, R. Opportunity in early education: Improving teacher-child interactions and child outcomes. In: Reynolds, A.; Rolnick, A.; Englund, M.; Temple, J., editors. Childhood Programs and Practices in the First Decade of Life: A Human Capital Integration. Cambridge University Press; New York: 2010. p. 243-265.

McCartney K, Burchinal MR, Bub KL. Best practices in quantitative methods for developmentalists. Monographs of the Society for Research in Child Development (SRCD). 2006; 71(3):1–145.

McCartney K, Burchinal M, Clarke-Stewart A, Bub K, Owen M, Belsky J, NICHD Early Child Care Research Network. Testing a series of causal propositions relating time in child care to children's externalizing behavior. Developmental Psychology. 2010; 46(1):1–17. [PubMed: 20053002]

McDermott PA, Fantuzzo JW, Waterman C, Angelo LE, Warley HP, Gadsden VL, Zhang X. Measuring preschool cognitive growth while it's still happening: The Learning Express. Journal of School Psychology. 2009; 47:337–366. [PubMed: 19712780]

Najarian, M.; Snow, K.; Lennon, J.; Kinsey, S. Early Childhood Longitudinal Study, Birth Cohort (ECLS-B), Preschool–Kindergarten 2007 Psychometric Report (NCES 2010-009). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education; Washington, DC: 2010.

NICHD Early Child Care Research Network. Does amount of time spent in child care predict socioemotional adjustment during the transition to kindergarten? Child Development. 2003; 74:976–1005. [PubMed: 12938694]

NICHD Early Child Care Research Network. Early child care and children's development in the primary grades: Follow-up results from the NICHD Study of Early Child Care. American Educational Research Journal. 2005; 42:537–570.

Nord, C.; Edwards, B.; Hilpert, R.; Branden, L.; Andreassen, C.; Elmore, A.; West, J. Early Childhood Longitudinal Study, Birth Cohort (ECLS-B), User's Manual for the ECLS-B Nine-Month Restricted-Use Data File and Electronic Code Book. National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education; Washington, DC: 2004.

Office of Head Start. Head Start Program Fact Sheet. Department of Health and Human Service; Washington, DC: 2010. U.S.from http://www.acf.hhs.gov/programs/ohs/about/fy2010.html [Retrieved July 28, 2011]

Paulson JF, Keefe HA, Leiferman JA. Early parental depression and child language development. The Journal of Child Psychology and Psychiatry. 2009; 50(3):254–262.

Radloff LS. The CES - D scale: A self-report depression scale for research in the general population. Journal of Applied Psychological Measurement. 1977; 1:385–401.

Rigby E, Ryan RM, Brooks-Gunn J. Child care quality in different state policy contexts. Journal of Policy Analysis and Management. 2007; 26:887–907.

Rodriguez ET, Tamis-LeMonda CS, Spellmann ME, Pan BA, Raikes H, Lugo-Gil J, Luze G. The formative role of home literacy experiences across the first three years of life in children from low-income families. Journal of Applied Developmental Psychology. 2009; 30:677–694.

Rosenbaum PR, Rubin DB. Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. American Statistician. 1985; 39(1):33–38.

Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. Biometrika. 1983; 70(1):41–55.

Royston P. Multiple imputation of missing values: Update of ICE. The Stata Journal. 2005; 5(4):527–536.

Rutter, M. Resilience reconsidered: Conceptual considerations, empirical findings, and policy implications. In: Shonkoff, JP.; Meisels, SJ., editors. Handbook of early childhood intervention. 2nd ed.. Cambridge University Press; New York: 2000. p. 651-682.

Rubin, D. Multiple imputation for nonresponse in surveys. John Wiley & Sons; New York: 1987.

Schnur ES, Brooks-Gunn J, Shipman V. Who attends programs serving poor children?: The case of Head Start attendees and non-attendees. Journal of Applied Developmental Psychology. 1992; 13:405–421.

Shonkoff, JP.; Phillips, DA., editors. From Neurons to Neighborhoods: The Science of Early Childhood Development. National Academy Press; Washington, DC: 2000.

Snow, K.; Derecho, A.; Wheeless, S.; Lennon, J.; Rosen, J.; Rogers, J.; Einaudi, P. Early Childhood Longitudinal Study, Birth Cohort (ECLS-B), kindergarten 2006 and 2007 data file user's manual (2010-010). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education; Washington, DC: 2009.

Taylor HH. Curriculum in Head Start. National Head Start Bulletin. 2000; 67:1–36.

The Infant Health and Development Program. Enhancing the outcomes of low-birth-weight, premature infants: A multisite, randomized trial. JAMA. 1990; 263:3035–3042. [PubMed: 2188023]

U.S. Department of Health and Human Services, Administration for Children and Families. Head Start Impact Study: First Year Findings. Washington, DC: May. 2005

U.S. Department of Health and Human Services, Administration for Children and Families. Head Start Impact Study: Final Report. Washington, DC: Jan. 2010

Waldfogel, J. What children need. Harvard University Press; Cambridge, MA: 2006.

Waterman C, McDermott PA, Fantuzzo JW, Gadsden VL. The matter of assessor variance in early childhood education-Or whose score is it anyway? Early Childhood Research Quarterly. 2012; 27:46–54.

Wen X, Leow C, Hahs-Vaughn DL, Korfmacher J, Marcus SM. Are two years better than one year? A propensity score analysis of the impact of Head Start program duration on children's school performance in kindergarten. Early Childhood Research Quarterly. in press.

Zhai F, Brooks-Gunn J, Waldfogel J. Head Start and urban children's school readiness: A birth cohort study in 18 cities. Developmental Psychology. 2011; 47:134–152. [PubMed: 21244155]

Table 1 Descriptive Statistics by Head Start and Other Specific Care Arrangements

	HS ($n \approx 1{,}150$)		Pre-K ($n \approx 1{,}050$)		Other center-based care ($n \approx 2{,}700$)		Other non-parental care ($n \approx 650$)		Parental care ($n \approx 1{,}350$)	
Variable	***M***	***SD***	***M***	***SD***	***M***	***SD***	***M***	***SD***	***M***	***SD***
School readiness outcomes at kindergarten										
Cognitive outcomes										
Expressive language	−0.17	1.06	0.14**	0.86	0.23**	0.86	0.04**	1.01	−0.08	1.07
Early readings	−0.31	0.91	0.17**	0.93	0.21**	0.91	−0.19	0.90	−0.18*	0.97
Mathematics	−0.31	0.88	0.13**	0.92	0.28**	0.91	−0.10**	0.93	−0.15**	0.97
Socio-emotional outcomes										
Conduct problems	0.18	0.92	0.05	0.87	−0.02**	0.83	−0.02**	0.77	−0.07**	0.85
Hyperactivity/Inattention	0.19	0.88	0.02**	0.85	−0.09**	0.81	0.01*	0.83	−0.09**	0.83
Pro-social behaviors	−0.03	0.70	0.08*	0.66	0.11**	0.68	−0.05	0.66	0.02	0.67
Child care arrangements										
9-months										
Parental care	0.50	0.50	0.47	0.50	0.45	0.50	0.33**	0.47	0.69**	0.46
Relative care	0.31	0.46	0.24*	0.43	0.23**	0.42	0.42**	0.49	0.19**	0.39
Non-relative care	0.12	0.33	0.17*	0.38	0.20**	0.40	0.22**	0.41	0.09	0.29
Center-based care	0.07	0.26	0.11*	0.32	0.12**	0.32	0.04	0.19	0.03**	0.18
2-years										
Parental care	0.51	0.50	0.46	0.50	0.45**	0.50	0.32**	0.47	0.75**	0.44
Relative care	0.24	0.43	0.19**	0.39	0.15**	0.35	0.36**	0.48	0.11**	0.32
Non-relative care	0.11	0.31	0.16*	0.36	0.19**	0.39	0.25**	0.43	0.08	0.27
Center-based care	0.14	0.34	0.20**	0.40	0.22**	0.41	0.07**	0.26	0.06**	0.24
Child characteristics										
Boys	0.50	0.50	0.51	0.50	0.50	0.50	0.55	0.50	0.49	0.50
Age in months at 9-months	10.41	1.90	10.63	1.93	10.45	1.99	10.31	1.83	10.36	1.86
Race/ethnicity at 9-months										

Variable	HS ($n \approx 1,150$)		Pre-K ($n \approx 1,050$)		Other center-based care ($n \approx 2,700$)		Other non-parental care ($n \approx 650$)		Parental care ($n \approx 1,350$)	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Non-Hispanic White	0.24	0.43	0.54**	0.50	0.67**	0.47	0.42**	0.49	0.45**	0.50
Non-Hispanic Black	0.29	0.45	0.17**	0.37	0.09**	0.29	0.16**	0.37	0.12**	0.33
Hispanic	0.39	0.49	0.22**	0.41	0.15**	0.36	0.35	0.48	0.36	0.48
Non-Hispanic Asian	0.01	0.11	0.04**	0.19	0.04**	0.19	0.02	0.14	0.02	0.15
Other	0.06	0.24	0.04	0.19	0.05	0.21	0.05	0.22	0.05	0.21
Low birth weight at birth	0.08	0.26	0.07	0.25	0.07	0.25	0.07	0.26	0.07	0.26
Prematurity at birth	0.14	0.35	0.12	0.33	0.11	0.31	0.12	0.33	0.12	0.32
Multiple birth at birth	0.04	0.19	0.04	0.18	0.05	0.21	0.03	0.17	0.03	0.18
Breast-feeding at 9-months										
None	0.41	0.49	0.27**	0.44	0.22**	0.42	0.35	0.48	0.32**	0.46
<3 months	0.24	0.43	0.22	0.41	0.22	0.42	0.26	0.44	0.21	0.41
3~6 months	0.17	0.38	0.24**	0.43	0.20	0.40	0.21	0.41	0.19	0.39
7 months plus	0.17	0.38	0.28**	0.45	0.36**	0.48	0.18	0.38	0.28**	0.45
Number of siblings at 9-months										
No sibling	0.42	0.49	0.39	0.49	0.44	0.50	0.40	0.49	0.34**	0.47
One	0.27	0.44	0.38**	0.49	0.36**	0.48	0.32	0.47	0.33*	0.47
Two or more	0.31	0.46	0.23**	0.42	0.20**	0.40	0.28	0.45	0.33	0.47
Health status at 2-years										
Poor/Fair	0.04	0.21	0.02*	0.15	0.01**	0.11	0.03	0.16	0.03*	0.16
Good	0.12	0.32	0.08*	0.27	0.09	0.29	0.14	0.34	0.10	0.29
Very good	0.31	0.46	0.26	0.44	0.25**	0.43	0.25	0.43	0.27	0.45
Excellent	0.53	0.50	0.64**	0.48	0.65**	0.48	0.59	0.49	0.61**	0.49
Pre-treatment scores at 2-years										
BSF (Bayley Short Form)	47.53	9.19	50.24**	10.94	52.27**	9.86	48.30	8.81	49.36**	9.64
ITSC (Infant/Toddler Symptom Checklist)	9.56	4.31	8.85**	4.26	8.65**	4.11	8.68**	4.27	8.44**	4.36
Maternal characteristics										

	HS ($n \approx 1,150$)		Pre-K ($n \approx 1,050$)		Other center-based care ($n \approx 2,700$)		Other non-parental care ($n \approx 650$)		Parental care ($n \approx 1,350$)	
Variable	***M***	***SD***	***M***	***SD***	***M***	***SD***	***M***	***SD***	***M***	***SD***
Age at birth	24.95	5.96	28.25**	5.90	29.23**	5.89	26.51**	6.13	26.66**	6.07
Mom lived with bio-mother until 16	0.82	0.39	0.87*	0.34	0.88**	0.32	0.84	0.36	0.84	0.36
Mom lived with bio-father until 16	0.50	0.50	0.64**	0.48	0.69**	0.46	0.60**	0.49	0.60**	0.49
Married at birth	0.47	0.50	0.72**	0.45	0.80**	0.40	0.59**	0.49	0.68**	0.47
Depressed at 9-months	0.52	0.51	0.40**	0.45	0.36**	0.41	0.45*	0.53	0.46	0.50
Employment status at 2-years										
Not working	0.53	0.50	0.40**	0.49	0.38**	0.48	0.29**	0.45	0.64**	0.48
Full-time	0.33	0.47	0.41**	0.49	0.38	0.48	0.53**	0.50	0.19**	0.39
Part-time	0.14	0.34	0.19*	0.40	0.25**	0.43	0.18	0.38	0.17	0.37
Foreign-born	0.31	0.46	0.21**	0.41	0.15**	0.35	0.20**	0.40	0.30	0.46
English is primary home language at 2-years	0.70	0.46	0.83**	0.37	0.88**	0.33	0.79**	0.41	0.71	0.45
Health status at 2-years										
Poor/Fair	0.12	0.33	0.07**	0.26	0.04**	0.19	0.10	0.30	0.12	0.33
Good	0.32	0.47	0.20**	0.40	0.19**	0.39	0.26	0.44	0.28	0.45
Very good	0.32	0.47	0.33	0.47	0.39**	0.49	0.32	0.47	0.30	0.46
Excellent	0.25	0.43	0.40**	0.49	0.38**	0.49	0.32*	0.47	0.30	0.46
Parenting behaviors										
KIDI at 9-months	5.73	2.13	7.00**	2.12	7.39**	2.10	6.50**	2.18	6.30**	2.23
Cognitively stimulating activities at 2-years	8.80	2.12	9.34**	2.03	9.70**	1.87	8.99	2.09	9.18**	2.06
Having sleeping routines at 2-years	0.77	0.42	0.87**	0.34	0.89**	0.31	0.83*	0.37	0.82*	0.39
Eating dinner together per week at 2-years	5.21	2.35	5.38	2.09	5.51**	2.05	5.33	2.37		1.99
No spanking at 2-years	0.60	0.49	0.63	0.48	0.68**	0.47	0.61	0.49	0.61	0.49
Family characteristics										
Parent's education at birth										
Below high school (0-11)	0.23	0.42	0.07**	0.26	0.05**	0.22	0.15**	0.36	0.20	0.40

	HS ($n \approx 1,150$)		Pre-K ($n \approx 1,050$)		Other center-based care ($n \approx 2,700$)		Other non-parental care ($n \approx 650$)		Parental care ($n \approx 1,350$)	
Variable	***M***	***SD***	***M***	***SD***	***M***	***SD***	***M***	***SD***	***M***	***SD***
High school (12)	0.41	0.49	0.21**	0.41	0.13**	0.34	0.34*	0.47	0.30**	0.46
Some college (13-15)	0.29	0.46	0.32	0.47	0.29	0.45	0.30	0.46	0.31	0.46
Above college (16+)	0.06	0.24	0.40**	0.49	0.53**	0.50	0.21**	0.41	0.19**	0.39
Parental occupational prestige at 9-months	1.78	2.05	2.53**	2.49	2.84**	2.60	2.79**	2.31	1.38**	2.05
Family income at 9-months										
$0-$20000	0.48	0.50	0.19**	0.39	0.12**	0.33	0.24**	0.43	0.33**	0.47
$20001-$35000	0.32	0.47	0.18**	0.38	0.17**	0.37	0.31	0.46	0.32	0.47
$35001-$50000	0.16	0.37	0.37**	0.48	0.36**	0.48	0.30**	0.46	0.27**	0.44
$50001+	0.04	0.20	0.27**	0.44	0.35**	0.48	0.15**	0.35	0.07	0.26
Family income at 2-years										
$0-$20000	0.46	0.50	0.17**	0.38	0.13**	0.34	0.24**	0.43	0.32**	0.47
$20001-$35000	0.31	0.46	0.19**	0.39	0.14**	0.34	0.29	0.46	0.29	0.45
$35001-$50000	0.19	0.39	0.36**	0.48	0.35**	0.48	0.31**	0.46	0.31**	0.46
$50001+	0.04	0.20	0.28**	0.45	0.39**	0.49	0.16**	0.37	0.08	0.27
Lived in urban area at 2-years	0.82	0.38	0.88**	0.32	0.87*	0.34	0.83	0.38	0.83	0.38
Region of country at 2-years										
Northeast	0.15	0.36	0.22**	0.42	0.19	0.39	0.10	0.31	0.09**	0.28
Midwest	0.18	0.39	0.13*	0.33	0.26**	0.44	0.27**	0.44	0.22	0.41
South	0.41	0.49	0.48*	0.50	0.31**	0.46	0.36	0.48	0.40	0.49
West	0.26	0.44	0.17**	0.37	0.24	0.43	0.27	0.44	0.30	0.46
Mom or child received WIC at 2-years	0.71	0.45	0.33**	0.47	0.24**	0.43	0.43**	0.49	0.54**	0.50
Family members received FS at 2-years	0.41	0.49	0.17**	0.38	0.11**	0.32	0.21**	0.41	0.26**	0.44
Family members received TANF at 2-years	0.14	0.35	0.05**	0.22	0.04**	0.20	0.08**	0.26	0.08**	0.27

Note. Sample sizes were rounded to the nearest 50, due to IES reporting rules. All numbers were weighted using 'W1R0' for the 9-month wave and 'W2R0' for the 2-year wave. Two-tailed *t* tests were conducted to compare means between HS participants and other children who belonged to each specific type of care arrangement before propensity score matching. The outcome measures were standardized to have a mean of 0 and a standard deviation of 1. WIC = Special Supplemental Nutrition Program for Women, Infants, and Children; FS = Food Stamp Program; TANF = Temporary Assistance for Needy Families.

* $p < .05$.

** $p < .01$.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Table 2

Head Start Participation and School Readiness Outcomes Compared with Other Specific Care Arrangements

	Early reading		Mathematics		Conduct problems	
Reference care	***M***	***(SE)***	***M***	***(SE)***	***M***	**(SE)**
Panel A. Pre-K						
Model 1	−0.38***	(0.04)	−0.28***	(0.04)	0.09	(0.05)
Model 2	−0.18***	(0.05)	−0.06	(0.04)	0.12*	(0.05)
Model 3	−0.18**	(0.06)	−0.06	(0.05)	0.09	(0.06)
Model 4	−0.19**	(0.06)	−0.07	(0.06)	0.08	(0.06)
Panel B. Other center-based care						
Model 1	−0.31***	(0.04)	−0.30***	(0.03)	0.14***	(0.04)
Model 2	−0.04	(0.04)	−0.01	(0.04)	0.03	(0.04)
Model 3	0.00	(0.05)	−0.00	(0.05)	0.05	(0.05)
Model 4	−0.01	(0.05)	−0.01	(0.04)	0.05	(0.05)
Panel C. Other non-parental care						
Model 1	0.01	(0.04)	−0.03	(0.04)	0.18**	(0.05)
Model 2	0.07	(0.05)	0.05	(0.05)	0.10	(0.05)
Model 3	0.12*	(0.06)	0.11*	(0.05)	0.07	(0.07)
Model 4	0.12*	(0.06)	0.10	(0.05)	0.05	(0.06)
Panel D. Parental care						
Model 1	0.04	(0.04)	0.04	(0.04)	0.19***	(0.04)
Model 2	0.07	(0.04)	0.08*	(0.04)	0.15**	(0.04)
Model 3	0.08*	(0.04)	0.09*	(0.04)	0.12*	(0.05)
Model 4	0.09*	(0.04)	0.08*	(0.04)	0.11*	(0.05)

Note. The results of expressive language, hyperactivity/inattention, and pro-social behaviors were not presented in Table 2 but available upon request. In Panel A, sample sizes for Model 1 (M1) and Model 2 (M2) were 2,200, 2,200, and 1,600; and those for Model 3 (M3) and Model 4 (M4) were 2,150, 2,150, and 1,550 in order of outcomes. In Panel B, sample sizes for M1 and M2 were 3,750, 3,750, and 2,850; and those for M3 and M4 were 3,800, 3,800, and 2,850 in order of outcomes. In Panel C, sample sizes for M1 and M2 were 1,750, 1,750, and 1,300; and those for M3 and M4 were 1,700, 1,700, and 1,250 in order of outcomes. In Panel D, sample sizes for M1 and M2 were 2,400, 2,400, and 1,750; and those for M3 and M4 were 2,400, 2,400, and 1,700 in order of outcomes. All reported sample sizes were rounded to the nearest 50, due to IES reporting rules. All models were weighted using 'WK1R0'. Model 1 included child characteristics; additional mother, parenting, and family characteristics were added in Model 2; Model 3 used propensity score matching, controlling for the same covariates with those of Model 2; and Model 4 was the same with Model 3 but further controlled for pre-treatment scores. The outcome measures were standardized to have a mean of 0 and a standard deviation of 1.

* $p < .05$.

** $p < .01$.

*** $p < .001$.

Table 3

Head Start Participation and School Readiness Outcomes by Subgroup

	Early reading		Mathematics		Conduct problems	
Reference care	***M***	***(SE)***	***M***	***(SE)***	***M***	***(SE)***
Panel A. Pre-K						
By initial cognitive ability						
Low cognitive ability (<= −0.50SD)	0.02	(0.10)	0.22*	(0.10)	−0.02	(0.13)
High cognitive ability (> −0.50SD)	−0.26***	(0.07)	−0.14*	(0.07)	0.12	(0.07)
By parental education						
High school or less	−0.16	(0.09)	−0.05	(0.08)	0.04	(0.09)
More than high school	−0.19*	(0.08)	−0.14	(0.08)	0.05	(0.08)
By number of hours in HS						
Part-time (20 hours or less)	−0.23**	(0.07)	−0.05	(0.07)	0.05	(0.07)
Full-time (more than 20 hours)	−0.12	(0.07)	−0.06	(0.06)	0.18*	(0.08)
Panel B. Other center-based care						
By initial cognitive ability						
Low cognitive ability (<= −0.50SD)	0.00	(0.09)	0.13	(0.09)	−0.04	(0.13)
High cognitive ability (> −0.50SD)	−0.07	(0.05)	−0.08	(0.05)	0.10	(0.06)
By parental education						
High school or less	0.04	(0.07)	0.02	(0.07)	−0.01	(0.08)
More than high school	−0.13*	(0.06)	−0.09	(0.06)	0.16*	(0.07)
By number of hours in HS						
Part-time (20 hours or less)	−0.04	(0.06)	−0.03	(0.06)	−0.02	(0.07)
Full-time (more than 20 hours)	0.01	(0.06)	−0.01	(0.06)	0.10	(0.08)
Panel C. Other non-parental care						
By initial cognitive ability						
Low cognitive ability (<= −0.50SD)	0.04	(0.11)	0.15	(0.12)	0.21	(0.13)
High cognitive ability (> −0.50SD)	0.10	(0.07)	0.07	(0.07)	0.01	(0.09)
By parental education						
High school or less	0.08	(0.08)	0.10	(0.07)	0.12	(0.09)

	Early reading		Mathematics		Conduct problems	
Reference care	***M***	***(SE)***	***M***	***(SE)***	***M***	***(SE)***
More than high school	0.11	(0.08)	0.03	(0.08)	0.13	(0.10)
By number of hours in HS						
Part-time (20 hours or less)	0.09	(0.06)	0.08	(0.06)	−0.02	(0.07)
Full-time (more than 20 hours)	0.19*	(0.07)	0.15*	(0.07)	0.12	(0.10)
Panel D. Parental care						
By initial cognitive ability						
Low cognitive ability (<= −0.50SD)	0.10	(0.08)	0.15	(0.09)	0.07	(0.09)
High cognitive ability (> −0.50SD)	0.06	(0.05)	0.08	(0.05)	0.09	(0.07)
By parental education						
High school or less	0.10*	(0.05)	0.12*	(0.06)	0.11	(0.07)
More than high school	0.05	(0.08)	0.03	(0.07)	0.17	(0.09)
By number of hours in HS						
Part-time (20 hours or less)	0.04	(0.05)	0.06	(0.05)	0.06	(0.06)
Full-time (more than 20 hours)	0.12*	(0.06)	0.11	(0.06)	0.19**	(0.07)

Note. The results of expressive language, hyperactivity/inattention, and pro-social behaviors were not presented but available upon request. In Panel A, sample sizes for the low cognitive ability group were 650, 600, and 450; those for the high cognitive ability group were 1,450, 1,450, and 1,050; those for the less-educated group were 950, 950, and 700; those for the more-educated group were 1,150, 1,150, and 850; those for the part-time group were 1,650, 1,650, and 1,200; and those for the full-time group were 1,600, 1,600, and 1,150 in order of outcomes. In Panel B, sample sizes for the low cognitive ability group were 1,100, 1,100, and 800; those for the high cognitive ability group were 2,600, 2,600, and 1,900; those for the less-educated group were 1,150, 1,150, and 850; those for the more-educated group were 2,650, 2,650, and 2,000; those for the part-time group were 3,250, 3,250, and 2,450; and those for the fulltime group were 3,200, 3,200, and 2,400 in order of outcomes. In Panel C, sample sizes for the low cognitive ability group were 450, 450, and 300; those for the high cognitive ability group were 1,150, 1,150, and 800; those for the less-educated group were 950, 950, and 650; those for the more-educated group were 750, 750, and 500; those for the part-time group were 1,150, 1,150, and 850; and those for the full-time group were 1,150, 1,150, and 800 in order of outcomes. In Panel D, sample sizes for the low cognitive ability group were 700, 700, and 700; those for the high cognitive ability group were 1,600, 1,600, and 1,150; those for the less-educated group were 1,350, 1,350, and 950; those for the more-educated group were 1,050, 1,050, and 750; those for the part-time group were 1,900, 1,900, and 1350; and those for the full-time group were 1,700, 1,700, and 1,300 in order of outcomes. All reported sample sizes were rounded to the nearest 50, due to IES reporting rules. All analyses were weighted using ‘WK1R0’ and based on Model 3 that used propensity score matching, controlling for all covariates. The outcome measures were standardized to have a mean of 0 and a standard deviation of 1.

* $p < .05$.

** $p < .01$.

*** $p < .001$.