



IRVINE NETWORK ON
INTERVENTIONS IN DEVELOPMENT

Distinctions without a difference? Preschool curricula and children's development

Working Paper
This version: May 2016

Jade Jenkins
University of California, Irvine

Anamarie Auger
RAND Corporation

Tutrang Nguyen and Winnie Yu
University of California, Irvine

Abstract

Public preschool programs often require the use of a research-based curriculum, yet limited research examines whether curricular decisions influence classroom processes and children's school readiness. This study uses three large samples of preschool children to examine differences in classroom quality and activities, in children's school readiness by curricular status (packaged curriculum vs. not), and across curricular packages. There were no significant differences in children's outcomes between classrooms that do and do not use a curriculum. Some significant differences existed in classroom activities across classrooms using different curricular packages; however, there exists substantial variability across classrooms implementing the same curricular package. Overall we find very little differences between preschool curricular packages for children's school readiness or classroom practices.

Keywords: Preschool Curricula, Early Childhood Education, Classroom Quality

Do preschool curricula promote child development? The vast majority of publicly-funded preschool programs require use of a “research-based curricula”. Head Start programs are mandated to use research-based “whole-child” curricula. Federally and state-sponsored quality rating and improvement systems (QRIS) incorporate curriculum into their rankings and consider the use of a developmentally appropriate, research-based curriculum as an indication of program quality (e.g., Auger, Karoly, & Schwartz, 2015). Tax dollars invested in funding public preschool programs are thereby also invested in curricula. These policies benefit publishers, but do they benefit preschool children?

Although most publishers claim their curricula are research-based, few describe the research on which the claim is based or how the curricula materials are explicitly linked to children’s development (Clements, 2007). The most popular curriculum used in Head Start and pre-kindergarten (pre-k) programs is the Creative Curriculum (46% of Head Start teachers according to Family and Child Experiences Survey (FACES); Hulseley et al., 2011), despite its rating by What Works Clearinghouse as “No Discernable Effectiveness” in promoting school readiness (U.S. Department of Education, 2013). The second most popular curriculum is HighScope (19% in FACES; Hulseley et al., 2001), whose only evidence comes from the Perry Preschool study; a small, intensive demonstration program conducted in the 1960s with counterfactual conditions that no longer apply to the current preschool population (i.e., children who did not attend center-based preschool; Belfield, Nores, Barnett, & Schweinhart, 2006; Duncan & Magnuson, 2013; Schweinhart, 2005).

In fact, we know very little about whether and how commonly used preschool curricula influence children’s school readiness. Also unknown is whether different curricular packages vary in terms of their prescribed classrooms activities or instructional practices. Prior research consists primarily of researcher-designed curricula studies implemented in highly controlled settings or using limited samples. Few studies have examined the relationship between preschool curricula, classroom activities, and children’s school readiness in authentic, business-as-usual preschool settings. Our study is a comprehensive examination of widely used preschool curricula and their associations with key activities in the preschool classroom environment that may relate to children’s academic and social-emotional outcomes. Using three large samples of low-income preschool children and their classroom experiences, we aim to understand how curricula packages drive learning activities when used in preschool programs operating at scale, and how in combination, these factors relate to children’s readiness for school.

Curriculum and Children’s Development

Curricula set goals for the knowledge and skills that children should acquire in an educational setting. They support educators’ plans for providing the day-to-day learning experiences to cultivate those skills through the provision of daily lesson plans, materials, and other pedagogical tools (Goffin & Wilson, 1994; Ritchie & Willer, 2008). Curricula differ across a number of dimensions such as philosophies, materials, the role of the teacher, pedagogy or modality (e.g., small or large group setting), classroom design, and child assessment. Preschool programs can choose their own curricula, but their choices are often constrained by a pre-

approved list developed by state agencies, accrediting bodies, or funding sources (Clifford & Crawford, 2009). Most programs, such as Head Start, require a curriculum that provides enriching experiences across the multiple domains of children's development (e.g., health, social-emotional, academic), known as "whole-child" curricula. The whole-child approach is anchored in Piagetian theory, which emphasizes child-centered active learning cultivated through strategic arrangement of the classroom environment (DeVries & Kohlberg, 1987; Piaget, 1976; Weikart & Schweinhart, 1987), and sociocultural theory, where the teacher provides supportive and responsive interactions with children (Vygotsky, 1978). Whole-child curricula purport to emphasize critical thinking and problem solving skills by providing open-ended learning opportunities and simultaneously cultivating the inter-related domains of children's development (Diamond, 2010; Elkind, 2007; Zigler & Bishop-Josef, 2006).

The curricula most widely used in Head Start and state pre-k programs are Creative Curriculum and HighScope, and packages like Scholastic and High Reach curricula are other common alternatives (Clifford et al., 2005; Hulseley et al., 2011; Phillips, Gormley, & Lowenstein, 2009). Despite their popularity, little empirical support exists for HighScope, none exists for Creative Curriculum, and neither curriculum has demonstrated effectiveness based on rigorous standards (Belfield et al., 2006; Preschool Curriculum Evaluation Research Consortium, 2008; Schweinhart, 2005; U.S. Department of Education, 2013).

The dearth of evidence supporting Creative Curriculum and HighScope is not unique to these curricula. Most recently, the National Center on Quality Teaching and Learning of the Office of Head Start (2014) released the "Preschool Curriculum Consumer Report", the first of its kind, which reviewed the most commonly used curricula in Head Start programs nationwide and provided ratings for each on a set of 13 criteria. One criterion is "Curriculum is Evidence-Based". Of the 14 curricula reviewed in the report, 7 had "no evidence", 5 had "minimal evidence", one had "some evidence", and only one was rated to have "solid, high-quality evidence" (Opening a World of Learning) with demonstrated effects on child outcomes. One of the first Institute for Education Sciences funded research projects was the Preschool Curriculum Evaluation Research Study Initiative, a large multi-site, random assignment experimental study of 14 different preschool curricula, of which only two curricula were found to be effective at promoting children's school readiness, one of which was a math-focused curricular add-on (2008). In other studies, well-implemented, small, randomized control trials of curricula that are less commonly used or are content-specific (i.e., target literacy or math skills) show positive impacts on skills targeted in the curricular materials (Bierman et al., 2008; Clements & Sarama, 2008; Diamond, Barnett, Thomas, & Munro, 2007; Fantuzzo, Gadsden, & McDermott, 2011; Morris et al., 2014).

Curricula and Early Childhood Education Policy

A surfeit of research shows that high-quality preschool *can* promote children's cognitive and physical development, particularly for low-income children (Barnett, 2011; Duncan & Magnuson, 2013; Gormley, Phillips, & Gayer, 2008). Yet the tremendous variability in preschool quality and effectiveness both within and between different types of programs (e.g.,

Head Start and pre-k) and within and between states reveal how little is known about precisely *what* makes preschool effective (Bloom & Weiland, 2015; Dotterer, Burchinal, Bryant, Early, & Pianta, 2009; Jenkins, 2014; Jenkins, Farkas, Duncan, Burchinal, & Vandell, 2016; Karoly, Ghosh-Dastidar, Zellman, Perlman, & Fernyhough, 2008; Walters, 2014).

Federal, state, and local policy efforts have used two main levers to improve the effectiveness of preschool programs: 1) funding professional development and teacher training in public preschool programs, and 2) creating quality improvement, licensing and monitoring systems (e.g., QRIS). While often overlooked, curricular requirements and curricula use are embedded in these and other policies that govern state and federal early care and education systems. Early learning programs (e.g., Head Start) prescribe that programs use curricula, curricula prescribe specific classroom activities using varying pedagogical approaches, and within this learning environment, preschool children learn and develop. Thus, curricula guidance or restrictions may be an efficient mechanism through which policy may improve the quality and effectiveness of preschool programs. By understanding how different curricula influence what happens in preschool classrooms and the key classroom activities that influence children's development, the early childhood education field can gain a deeper understanding of factors that may make preschool effective for low-income children. Our goal is to identify the specific stages of this process and the associated curricular packages that may be the most promising at promoting children's school readiness skills. We present an illustration of this theoretical model in Figure 1.

Another critical policy consideration is that curricula are a significant investment for preschool programs. In Table 1 we present the approximate costs per classroom for commonly used curricula, which range between \$1125-\$4190 per classroom. Not included in these estimates are the additional professional development activities often strongly recommended by publishers to implement the curricula with fidelity, and the costs of supplemental materials. The Head Start program alone has over 50,000 classrooms, making the costs of such policies nontrivial (Office of Head Start, 2010). Given the wide array of curricular choices available, the government expenditures for the curricula required in public preschool, and our insufficient understanding of whether whole-child curricula promote children's school readiness, a comprehensive study of preschool curricula is badly needed.

Present Study

Our study uses three samples of preschool children to examine how commonly used published preschool curricula relate to classroom activities and quality, and subsequently to children's academic and social-emotional development in classrooms where teachers report using a curriculum. Specifically, our research questions (RQ) are:

1. To what extent does implementing a published "whole-child" preschool curriculum relate to classroom activities relative to classrooms without a published curriculum in use?
2. Do children's academic and social-emotional school readiness differ by classroom curricular status (i.e., published curriculum in use versus not)?
3. To what extent do classroom activities vary across classrooms using different whole-child

curricula?

4. Do children's academic and social-emotional school readiness differ by classroom curricula?

Method

Data

Our study uses secondary data from three studies of children in preschool settings between the 2001 and 2004 school years: the Preschool Curriculum Evaluation Research Study (PCER), the National Center for Early Learning and Development Multi-State Study of Pre-Kindergarten (NCEDL), and the Head Start Impact Study (HSIS). Each dataset contains information on curricula, classroom activities, and child academic and social-emotional outcomes. In all three studies, data collection took place in center-based preschool environments and the child participants were majority low-income and were ethnically and racially diverse. We describe each study's sample and measures in the following sections.

Samples.

PCER. Beginning in 2003, 12 grantees across the country were funded to study the effect of preschool curricula on children's academic and social-emotional outcomes in the PCER study. Each grantee selected their study curricula for a total of 14 different curricula tested in 18 different locations. Mathematica Policy Research and Research Triangle Institute assisted with the evaluation to ensure consistent data collection at each site, but each grantee was in charge of its own evaluation. Individual grantees were responsible for recruiting preschool centers to participate in the study. At each grantee site, either classrooms within preschool centers or entire centers themselves were randomly assigned to a treatment (experimental curriculum) or control condition. For feasibility and to preclude cross contamination across classrooms, most research sites assigned only one curriculum to each preschool center. Baseline data on children, parents, and preschools were collected in the fall of 2003, with post-treatment data collected in the spring of 2004. Approximately 2,900 children in 320 preschool classrooms participated in the study. The subsample of PCER most relevant to our study are the grantee sites and classrooms that used one of our focal whole-child curricula—HighScope, Creative Curriculum, DLM Express—and those classrooms with no published curriculum in use (N=1,450 children). The data include children who were either in Head Start, private child care, or public preschool. For more information on the study, see PCER Final Report (2008).

NCEDL. This study is comprised of two stratified random samples of children within preschool programs across 11 states. States were purposely selected if they had large numbers of children enrolled in pre-existing public pre-k programs. The sample for the Multi-State Study of Pre-Kindergarten includes 6 states (California, Illinois, Georgia, Kentucky, New York, and Ohio). The follow-up study, the State-Wide Early Education Programs Study was not included in our analyses because the dataset did not include curriculum indicators. Preschool programs were randomly sampled within states, and 29% were Head Start programs. One classroom was then randomly sampled within each program, and 94% of classroom teachers agreed to participate. Of

the selected classrooms, approximately 60% of parents gave consent for their child to participate, and from this subsample four children were randomly selected to participate (N=1,015). Forty preschool programs were selected in each state for a total of 245 classrooms. Child assessment data were collected during the fall and spring of the 2001-2002 preschool year. For more information, see Early et al. (2005).

HSIS. The HSIS is a nationally representative study of Head Start participants and a group of comparable non-participants from 23 states that were sampled using a complex multi-stage stratified design. Head Start programs (grantees) were divided into geographic clusters and were then stratified based on program characteristics, with three grantees or delegate agencies randomly selected from each cluster. Within each delegate agency, Head Start centers were stratified in the same way as grantees, and were randomly selected. This resulted in 84 programs and delegate agencies with a total of 383 individual preschool centers. The full sample included newly entering 3- and 4-year-old Head Start applicants at randomly selected oversubscribed centers, where children were randomly assigned to receive an offer for Head Start. A total of 4,442 children were selected – 2,646 for Head Start and 1,796 for the control condition. Control group participants either found other available child care services (including Head Start in some cases) or the child was cared for at home. Study investigators (Westat) collected baseline surveys and child assessments during fall of the preschool year (2002), and post-treatment child assessments were collected at the end of Head Start in the spring. We restrict the sample for our study to those children who were randomly assigned to, and actually attended, a Head Start program because only under these conditions were classrooms required to have a curricular package in use. Control children in the HSIS were omitted from our study due to the extensive variation in counterfactual care conditions. For more information, see the HSIS Final Evaluation Report (Puma et al., 2012).

Measures.

Preschool curricula. Each dataset includes classrooms using one of our study's focal curricula, as outlined in Table 1. Additionally, both the NCEDL and PCER samples include preschool classrooms with no published curriculum in use. No published curriculum means that the classroom did not use a published or packaged curriculum but may have used a "locally developed" or a teacher-designed curriculum. Although we cannot know the exact content of these curricula or the curricula models on which they are based, we consider the no published curriculum and the locally or teacher developed curriculum designations to represent another common practice in early childhood education and thus important to include in our study. In the NCEDL and the HSIS studies, a category indicating "Other published curricula" represents those classrooms for whom we do not have specific curricular package information, or where there were too few classrooms using a specific whole-child curriculum and were collapsed into a single group for analysis.

We acknowledge that teachers may report "using" a curriculum when it may merely be present on their classroom bookshelves. However, we wish to highlight that the aim of our study is to understand the implications of policy-mandated curricula. As such, our data represent the

de facto, classroom environments for children who experienced different curricular choices with at scale business-as-usual implementation.

Classroom quality. Quality of care was measured with several instruments across the three studies. The *Early Childhood Environment Rating Scale-Revised* (ECERS-R; Harms, Clifford, & Cryer, 1998) is a widely used observer-rated measure of global classroom quality, specifically designed for use in classrooms serving children between 2.5 and 5 years of age, and was used in each study. Scores on the ECERS-R range from 1-7 with 1 indicating “inadequate” quality, 3 indicating “minimal” quality, 5 indicating “good” quality, and 7 indicating “excellent” quality. The scale’s authors report a total scale internal consistency of .92. To capture caregiver interactions, the HSIS and PCER studies used the *Arnett Caregiver Involvement Scale* (Arnett, 1989). This is an observational measure consisting of 26 items reflecting teacher sensitivity, harshness, and detachment that are rated on a 1-4 scale indicating how characteristic they are of the teacher, from not at all (1) to very much (4). Psychometric analyses suggest that the items load onto a single factor (Cronbach’s $\alpha = .93$).

Classroom learning activities. We used different instruments and data sources in each study to create aggregate measures of total classroom literacy and mathematics activities. Detailed lists of the individual items used, along with mean values by curricular package, are available in Appendix A1-A3.

The *Teacher Behavior Rating Scale* from the PCER study uses trained observers to rate the quality and quantity of academic activities present in a classroom (Landry et al., 2001). There are two content areas measured by the TBRS - math and literacy. Literacy is composed of five subdomains (written expression, print and letter knowledge, book reading, oral language, and phonological awareness). Quality of activities were rated from 0-3 (0 = activity not present; 3 = activity high quality). Quantity of activities was similarly rated from 0-3 (0 = activity not present; 3 = activity happened often or many times). We focus only on the quantity measures in our analyses, and this number was derived from taking the average of each of the activities that were rated. Cronbach’s α for the math scale is .94, and for the literacy scale, .87.

The *Emerging Academic Snapshot* (EAS) used in the NCEDL study is also an observer rated measure of children’s classroom engagement that captures children’s moment-to-moment activities (Ritchie, Howes, Kraft-Sayre, & Weiser, 2001). Observations were conducted over 1 or 2 days in the spring of the preschool year. The data collector observes each study child in 20-second interval “snapshots,” followed by a 40-second coding period. The other three study children in the sampled classroom are then coded before coming back to observe the first child again, and this is repeated for the entire observation period. Children are coded with one of six mutually exclusive activity settings in each snapshot (basics, free choice, individual time, meals, small group, and whole group). The activity is also coded for early academic content area (aesthetics, fine motor skills, gross motor skills, letter and sound, mathematics, oral language development, read to, science, social studies, and writing). For example, to obtain the proportion of the day spent in mathematics activities at the classroom level, coders took the average amount of time that each sample child was observed being engaged in mathematics activities divided by

the total observation time. The last coded component of each snapshot is the type of teacher–child interaction (routine, minimal, simple, elaborated, scaffolding, and didactic). Kappas range from .70 to .87.

End of year teacher surveys were used in the HSIS data to capture the different types of classroom activities. Teachers were asked how many times in the past week their class engaged in a specific literacy or math activity (shown in Appendix A3). We used the teacher-reported items on the type and frequency of classroom literacy and math activities, converted into times per month by taking the mean value of the answer category (e.g., never = 0; 1-2 times per week = 1.5) multiplied by 4, following Claessens, Engel, and Curran (2013). We then standardized this measure to have a mean of 0 and standard deviation of 1. Prior research indicates that teacher survey instruments are valid for assessing quantity of instruction, but not quality (Herman, Klein, & Abedi, 2000).

Child school readiness skills. Our analyses use multiple literacy, language, math, and social-emotional assessments that are considered valid and reliable, and are widely used within the field of child development. In each study, children were assessed at the beginning and end of their preschool year so that the baseline score can be used as a control variable in all analyses. Note that we do not include descriptions of PCER’s child school readiness measures because we are unable to estimate our child school readiness analytic models using those data (see Analyses below for detail).

Receptive language was measured by the *Peabody Picture Vocabulary Test* (PPVT; Dunn & Dunn, 1997) in each study, which focuses on children’s ability to successfully point to the picture that most closely represents the word spoken to them by the test administrator. Reliability for the PPVT ranges from .92 to .98. Children’s emergent literacy skills were also measured with the *Letter Word* subtest from the *Woodcock-Johnson Psycho-Educational Battery-Revised III* in the HSIS (WJ; Woodcock, McGrew, & Mather, 2001). In the Letter Word (LW) test, the child is initially asked to identify letters and as the test progresses in difficulty, children are asked to read and pronounce written words correctly. This assessment measures children’s ability to correctly recognize and sound out letters and sight words. Reliability is between .97 and .99 for preschool children. The HSIS also included the *WJ Spelling* subtest. The Spelling subtest requires children to trace letters, write letters in upper and lowercase, and to spell words, measuring early writing and spelling skills (Cronbach’s $\alpha = .90$).

Children’s general mathematical knowledge was assessed by the *WJ Applied Problems* subtest in both studies (Woodcock et al., 2001). The Applied Problems (AP) subtest examines the child’s ability to analyze and solve math problems. The reliability coefficient for the 3- to 5-year-old age group ranges from .92 to .94.

The *Teacher-Child Rating Scale* (TCRS; Hightower, 1986) was used to measure children’s social and emotional skills in the NCEDL study. This is a behavioral rating scale that assesses children’s social competence and problem behaviors. The Social Competence scale was computed as the mean of 20 items and had a Cronbach’s α of .95. The Problem Behavior scale was computed as the mean of 18 items and had a Cronbach’s α of .91. The HSIS study included

the 28-item *Behavior Problems Index* (Zill, 1990). This is a parent report of problem behaviors related to emotional status, school behavior, and interpersonal relationships, with items drawn from several other child behavior scales (e.g., CBCL). Items are rated on a 3-point scale. Zill reported 2-week test-retest reliability of .92.

Covariates. Each dataset contains several child and parent characteristics that are included as control variables in our analyses. These include gender of child, race of child, mother or primary caregiver educational level and age, and family income. Data on these characteristics were collected from parents via parent report during the preschool year. We also include children's baseline outcome assessments from the fall of the preschool year as covariates. In the NCEDL analyses we include an indicator for family poverty as a control variable, and in the HSIS analyses, an indicator for teen mother, due to the nature of these two samples. The classroom, teacher, and center covariates include teachers' education and years of experience, classroom-level aggregates of children's race, gender, and parental education, and whether the classroom is located in a public school or is a Head Start provider (PCER and NCEDL only). Because PCER was an experimental study, we control for classroom treatment status in our analyses.

Analyses

RQ 1 & 3: Descriptive analyses. A first-order question in the investigation of preschool curricula and children's school readiness is whether differences exist in the way children spend their time in preschool classrooms and in their school readiness skills by curricular status (i.e., published curriculum in use: yes or no; RQ 1), or by curricular package (e.g., HighScope compared with Creative Curriculum; RQ 3). First, we compare the available measures of classroom activities, quality, and other key classroom features (e.g., teachers' education) by curricular status in PCER and NCEDL using t-tests of means or z-tests of proportions. We also run OLS regressions with the classroom activities as the dependent variables and controlling for other classroom characteristics to further test for associations between curricular status and classroom activities. Because all Head Start classrooms are required to use a curriculum, we could not test for differences in classroom features by curricular status in the HSIS.

To examine the variance in classroom activities between different curricula packages (RQ 3), we conduct ANOVAs in each of the three datasets. We also regress classroom activities on indicators for curricular package with other classroom characteristics controlled.

RQ 2 & 4: Program and state fixed effects. Curricula are not randomly assigned to programs, centers, teachers, or children, and therefore we cannot causally determine whether a curriculum affects children's school readiness. Unobserved or unmeasured characteristics may be affiliated with both curricula and children's outcomes. Therefore, we test for associations between curriculum status (RQ2) and curricular package (RQ4) and child school readiness outcomes using two different types of fixed effects. Fixed effects is an econometric technique that removes from the estimate of interest any context-specific and time-invariant observable or unobservable characteristics that may influence both the choice of curriculum and children's outcomes. In essence, these models control for differences in a context shared by children. In

our study, this meant including separate intercepts (i.e., indicator variables) for each unique context (e.g., Head Start grantee).

The analysis most robust to bias from unobserved center and classroom characteristics comes from the HSIS data, where we are able to estimate a program (grantee) fixed effects model. This method takes advantage of differences in classroom curricula within the programs where families applied for, and were randomly assigned to receive, Head Start services at one of the centers operated by that program. In other words, this analysis allows us to compare the outcomes of children living in the same area who received Head Start services from the same grantee, reducing the possibility of omitted variables bias, but not eliminating it. The general form of this model is as follows:

$$\text{ChildOutcome}_{ij} = b_0 + a(\text{Curricula}_c) + x(\text{Child Controls}_i) + q(\text{Classroom Controls}_c) + p(\text{HS Program}_z) + e_{ijc} + u_{ij}$$

where ChildOutcome represents a child's (*i*) school readiness outcome (*j*; e.g., PPVT) at the end of the preschool year, *a* is a vector of curriculum indicator variables (or a curricular status indicator) which vary by classroom (*c*), *x* is a vector of child and family control variables which also includes children's baseline assessments, *p* is a vector of indicators for each (*z*) of the Head Start programs included in the study, *e* represents unaccounted for child, family, and classroom characteristics, and *u* reflects the remaining sources of variation in children's school readiness. The coefficients in *a* are our estimates of interest since they represent the differential associations between each preschool curriculum (or curricular status) and children's school readiness relative to the reference curriculum. Of the 84 programs in the HSIS, 62 (75%) had variation across classrooms in curricular package, with Creative Curriculum as the most common.

Since curricula are not randomly assigned, the interpretation of *a* must allow for the possibility that curricula will be picking up other classroom or center characteristics that are correlated with curricula. We attempt to minimize this problem by including appropriate teacher, classroom and center-level controls available, indicated by *q* (i.e., teacher's education, teacher's years of experience, and ECERS score). We adjust for the clustered sample design using Huber-White standard errors.

Although we are unable to estimate a similar program fixed effects model in the PCER and NCEDL because of the differences in sampling strategies and study designs, we estimate a state fixed effect model with the NCEDL dataset. This model compares children in preschool classrooms within the same state across classrooms using different curricula, with Creative Curriculum as the reference group. This model replaces *p* in the equation above with indicators for states in the NCEDL study. This is done to mitigate bias from state-level policies, regulations, and funding streams affecting preschool quality or curricular requirements.

Because no common curricular reference group exists across states or grantees in the PCER study, we are unable to test for differences in child outcomes by curricular package. There were also not enough states or grantees in the PCER sample with variation in curricular status (RQ 2) to test for differences in outcomes. In total, we examine relationships between curricula and children's outcomes in the HSIS and NCEDL samples only.

Results

RQ 1: To What Extent Does Implementing a Published “Whole-Child” Preschool Curriculum Relate to Classroom Activities Relative to Classrooms Without a Published Curriculum in Use?

We computed descriptive statistics and t-tests to assess whether having a curricular package in use makes a difference in terms of children’s preschool classroom experiences in the PCER and NCEDL samples, presented in Table 2. All Head Start classrooms use curricula, and therefore the HSIS data are omitted from this analysis.

The PCER results indicate that classrooms reporting use of a published curriculum have significantly ($p < .05$) more literacy and math activities and higher quality ratings on both the total ECERS score and Arnett scale relative to classrooms where the teacher reports using no formal or published curriculum. However, OLS regressions controlling for teachers’ education and years of experience and classroom-level aggregates of children’s race, gender, and parental education indicate that the differences are no longer significant (Appendix B1).

Our descriptive analyses in the NCEDL sample revealed that classrooms using a published curriculum score higher on the total ECERS score and in the Provisions for Learning ECERS factor compared with classrooms that were not using a curriculum. No significant differences emerge in the amount of classroom math and literacy learning activities by curricular status. These results were confirmed in state fixed effects regression models restricting the comparison of classrooms with and without a curriculum in use to those within the same state, controlling for teachers’ education and years of experience and classroom-level aggregates of children’s race, gender, and parental education (Appendix B1).

RQ 2: Do Children’s Academic and Social-emotional School Readiness Differ by Classroom Curricular Status?

State fixed effects models testing for differences in children’s school readiness by curriculum status in NCEDL are presented in Appendix B2. We find no significant differences in children’s math, literacy, or social skills depending on whether the classroom used a published curriculum. However, we find that teachers reported significantly fewer problem behaviors in classrooms where a curriculum was used.

RQ 3: To What Extent Do Classroom Activities Vary Across Classrooms using Different Whole-Child Curricula?

We present descriptive statistics and ANOVAs for each dataset in Table 3 to examine differences in the means and proportions of classroom activities and characteristics by curricular package. In all three samples there were significant differences across the curricular packages in both the quantity of math activities and overall classroom quality based on ECERS scores and the Arnett Caregiver Interaction Scale. Other significant differences by curricular package are teacher’s education (PCER, NCEDL, HSIS), length of school day and program location (NCEDL, HSIS), and teacher wages and Head Start status (NCEDL). Although we find significant differences between curricular packages on some measures in each of the three datasets, across the three datasets taken together, no clear distinction emerges for any curriculum

in terms of its allocation of literacy and math activities or its superior quality (e.g., HighScope did not have consistently higher ECERS ratings than Creative Curriculum).

Regressions of classroom activities on indicators for curricular package controlling for other classroom characteristics are presented in Appendix D. As a complement to the ANOVAs, this analysis allowed us to directly compare each curriculum with the reference category (Creative Curriculum), to see which—if any—was a top performer on classroom activities or quality. These results were very similar to the patterns in Table 3, with no clear front-runner across the three datasets. The most substantial differences came from the HSIS analyses, revealing that Scholastic classrooms score significantly lower than Creative Curriculum classrooms on the ECERS (-0.40 SD), and significantly higher on literacy activities (0.20 SD).

RQ 4: Do Children’s Academic and Social-Emotional School Readiness Differ by Curricula?

We present the results for models examining differences in children’s school readiness outcomes by curricular package for the HSIS and NCEDL samples in Table 4. The reference group is Creative Curriculum in both datasets. All outcomes are in standard deviation (SD) units so that coefficients can be interpreted as effect sizes.

The HSIS program fixed effects models indicate that after controlling for Head Start program—and thus as many unobserved center-level factors as possible—children in Head Start classrooms using the Scholastic curriculum consistently outperform children in other classrooms operated by that program using the Creative Curriculum on each outcome except for behavior problems. The differences are also substantial in magnitude: we detect between 0.15-0.52 SD differences in children’s outcomes between Scholastic and Creative Curriculum classrooms. For the WJ Spelling subtest scores, children’s assessments were significantly lower in classrooms using Creative Curriculum compared with HighScope, Scholastic, and the “other curricular packages” set of classrooms. Children in classrooms using the HighScope curriculum also scored 0.2 SD higher on the WJ Applied Problems subtest compared with children in Creative Curriculum classrooms. Children in classrooms using High Reach scored significantly worse on PPVT scores relative to Creative Curriculum.

In state fixed effects models in the NCEDL dataset, we find that children in classrooms with no published curriculum in use have higher PPVT scores relative to Creative Curriculum classrooms. However, they also have higher teacher-reports of problem behavior, which parallels the RQ1 finding that classrooms with no curriculum had higher levels of problem behavior. Otherwise, there were no significant differences in children’s outcomes between classrooms using HighScope, other curricular packages (not specified), or no curriculum relative to the Creative Curriculum. The standard errors are quite large for these analyses, which limit our ability to draw inferences.

Because the descriptive analyses from NCEDL showed other differences in classrooms characteristics across curricular package, we ran additional regression models with the full set of control variables including whether the program was full-day, Head Start, and all teacher characteristics, to mitigate unobserved variables bias. The estimates remain the same (Appendix

E).

Variation in the Implementation of Curricula

Although we found some differences by curriculum on children's outcomes, classroom quality, and the quantity of classroom math activities, our findings were not consistent across the three samples. These inconsistencies led us to question whether classrooms using the same curriculum were highly variable with respect to their observed classroom processes. In other words, we asked, "are the implementation of activities and overall classroom quality consistent amongst classrooms using the same curricular package?" Indeed, an important aspect of curricula with respect to public preschool policy has to do with implementation. In theory, the curriculum drives classroom activities. Thus we hypothesized that classrooms where teachers reported using the same curriculum would be comparable with respect to quantity and type activities and, perhaps, overall quality.

To examine the variability in classroom experiences across classrooms implementing the *same* curriculum, we present histograms of ECERS scores and the frequency of literacy and math activities for the two most popular curricula, Creative Curriculum and HighScope. Figure 2a shows the distributions of ECERS scores, and math and literacy activities in NCEDL (left) and PCER (right) for Creative Curriculum classrooms, and Figure 2b shows the same distributions for HighScope classrooms. Each measure is in its original scale (i.e., not standardized). Overlaid on these graphs are the distributions for classrooms that *do not* report using a published curriculum, for comparison.

There are two striking and consistent features of these graphs: 1) The distributions of activities and quality across classrooms using Creative Curriculum or HighScope are nearly identical to the distributions from classrooms who do not report using any published curriculum; and 2) Classrooms using the same curriculum vary widely with respect to their overall quality and learning activities. Furthermore, these illustrations help to explain our largely null findings thus far; substantial variation in the learning experiences *within* the population of classrooms that report using HighScope or Creative Curriculum would not likely lead to systematic differences in children's outcomes *across* curricular packages.

We could not do the comparison overlay in the HSIS graphs because we only use Head Start treatment children in our study, and all classrooms are required to use a published curriculum. (The measurement scales in NCEDL and PCER are different from the HSIS and so overlaying those distributions on the HSIS classrooms would be difficult to interpret.) Histograms from the HSIS sample without the "no published curriculum" overlays are available in Appendix F. Similarly, these figures indicate substantial dispersion or variation in the distribution of activities and quality.

Discussion

Our study comprehensively examined the role of curricula in center-based preschool environments and their relation to children's academic and social-emotional development in three different preschool studies. The three samples we used capture the authentic preschool experiences of a diverse set of low-income children attending public and private pre-k programs

and of a nationally representative sample of children attending Head Start programs. Specifically, our research questions were: 1) To what extent does implementing a published “whole-child” curriculum relate to classroom activities without a published curriculum in use? 2) Do children’s academic and social-emotional school readiness differ by classroom curricular status? 3) To what extent do classroom activities vary across classrooms using different whole-child curricula? 4) Do children’s academic and social-emotional school readiness differ by classroom curricula? In two datasets our measure of curricular implementation was captured by observer-rated classroom activities (PCER, NCEDL), and in one, teachers’ reports of classroom activities (HSIS). We do not equate implementation with fidelity, and consider what we observe as the de facto use of curriculum in at scale preschool programs. Our goal was simply to describe the observed patterns in extant data and glean as much information as possible given the developmental and policy relevance of our research questions.

We did not find consistent evidence that having a classroom curriculum was associated with higher quality scores on the ECERS or the Arnett Caregiver Interaction Scale, or more frequent academic activities. One curriculum used in the HSIS—Scholastic—stood out by being the most related to math activities in the classroom and children’s academic school readiness. However, Scholastic was not included in the PCER or NCEDL studies so we could not replicate these analyses in other datasets. Still, this finding may be particularly promising for preschool programs. It is also worth noting that Scholastic had significantly lower quality scores than Creative Curriculum (Appendix D). Future research should examine the Scholastic preschool curriculum to understand the curricular elements that may positively influence classroom activities and children’s outcomes.

Taken together, the findings from our study indicate that there are few distinguishing characteristics about the curricula most commonly used in today’s preschool programs. When a teacher reports using a published curricular package, the features of their classroom are not distinguishable from classrooms where teachers report using no published curriculum. For classrooms that reported using HighScope and Creative Curriculum, the two most popular curricula, classroom literacy and math activities and ECERS scores varied as widely within the population of classrooms using each package as it did across the population of preschool classrooms where teachers report not using any published curriculum.

In fairness, we must consider the perspective of curricula publishers and developers in interpreting our results. Our data represent the authentic educational environments of low-income preschoolers. Thus, they do not necessarily represent classroom experiences when curricula are implemented with high fidelity (which we cannot measure in this study), and with developer-specified professional development; in other words, our analyses do not represent tests of curricular efficacy, but represent the business-as-usual experiences of children in public and private preschool programs. Indeed, teachers may report “using” a curriculum that they only reference on occasion (or not at all). The aim of our study was to understand the implications of policy-mandated curricula and thus our data represent the de facto educational environments for children attending preschool during 2001-2004. The policy requirements would necessarily need

to change and include greater professional development and other supports to implement curricular packages with high fidelity at scale. Still, *none of the curricula under study have evidence of efficacy under ideal conditions*, so this criticism on its own falls short of how we need to think carefully about curricular choices in public preschool programs for low-income children. An important consideration is whether the variation in curricular implementation found in our study is what one would expect from a sample of publicly funded preschool classrooms, a priori. Future research might consider how much variation in curricular implementation is expected, and the role of professional development to ensure consistently high-quality classroom experiences for children in preschool programs.

Our findings, while primarily descriptive in nature, beg an extremely important and policy relevant question: what do current curricular investments in early childhood policy yield for children's development and well-being? We do not find any evidence to support Creative Curriculum as the dominant choice for use in public preschool programs, corroborating the What Works Clearinghouse rating of "No Evidence", nor do we find support for other curricula with the exception of Scholastic. We caution against interpreting these results as causal and instead suggest that they be a starting point for future research and policy discussions.

A strength of our paper is replication across three different samples of preschoolers. However, this also means that the unique components of each dataset restricted us from universally conducting the same analyses. We also recognize that our measures of curricular implementation and quality are limited and do not capture the full set of preschool classroom experiences shaped by curricular packages. Note that many of the PCER classrooms implementing randomly assigned curricula had study administrators in the classroom providing professional development to help teachers implement the curriculum. Although we conducted additional analyses to control for treatment status with the PCER data, these additional supports may not generalize to other preschool classrooms.

Future research should consider the specific activities that are associated with children's development along with a content examination of curricula to examine which packages promote the most beneficial activities. In addition, content analyses could illuminate how well-aligned curricula are with early learning standards set by states and national organization such as NAEYC, akin to those conducted with elementary and secondary school curricula (Polikoff, 2015; Porter, 2002; Schmidt, Wang, & McKnight, 2005). Specifically, in our work, we are beginning to examine the content of the curricular packages most commonly used in public preschool programs by conducting a content analysis of the teacher's manuals for different whole-child curricula. Developing this evidence base will provide a deeper understanding of factors that may make preschool effective for low-income children.

Acknowledgements

We are grateful to the Institute of Education Sciences (IES) for supporting this work through grant [redacted] awarded to [redacted]. Research reported in this publication was also supported by the Eunice Kennedy Shriver National Institute of Child Health & Human Development of the National Institutes of Health under Award Number [redacted]. The content

is solely the responsibility of the authors and does not necessarily represent the official views of IES, the U.S. Department of Education, or the National Institutes of Health. We would also like to thank [redacted] for helpful comments on prior drafts.

References

- Auger, A., Karoly, L., & Schwartz, H. L. (2015). *Evaluation of Delaware Stars for Early Success: Year 2 Report*. RAND Corporation. RAND Corporation: Santa Monica, CA.
- Barnett, W. S. (2011). Effectiveness of Early Educational Intervention. *Science*, 333(6045), 975-978. doi:10.1126/science.1204534
- Belfield, C. R., Nores, M., Barnett, W. S., & Schweinhart, L. J. (2006). The High/Scope Perry Preschool Program. *Journal of Human Resources*, XLI(1), 162-190. doi:10.3368/jhr.XLI.1.162
- Bierman, K. L., Domitrovich, C. E., Nix, R. L., Gest, S. D., Welsh, J. A., Greenberg, M. T., . . . Gill, S. (2008). Promoting Academic and Social-Emotional School Readiness: The Head Start REDI Program. *Child Development*, 79(6), 1802-1817. doi:10.1111/j.1467-8624.2008.01227.x
- Bloom, H. S., & Weiland, C. (2015, Nov 7). *Quantifying Variation in Head Start Effects on Young Children's Cognitive and Socio-emotional Skills using data from the National Head Start Impact Study*. Working Paper. New York, NY.
- Claessens, A., Engel, M., & Curran, F. C. (2013). Academic Content, Student Learning, and the Persistence of Preschool Effects. *American Educational Research Journal*. doi:10.3102/0002831213513634
- Clements, D., & Sarama, J. (2008). Experimental Evaluation of the Effects of a Research-Based Preschool Mathematics Curriculum. *American Educational Research Journal*, 45(2), 443-494. doi:10.3102/0002831207312908
- Clifford, R. M., Barbarin, O. A., Chang, F., Early, D. M., Bryant, D., Howes, C., . . . Pianta, R. (2005). What is Pre-Kindergarten? Characteristics of Public Pre-Kindergarten Programs. *Applied Developmental Science*, 9(3), 126-143.
- Clifford, R. M., & Crawford, G. M. (Eds.). (2009). *Beginning School: U.S. policies in international perspective*. New York: Teachers College Press.
- DeVries, R., & Kohlberg, L. (1987). *Programs of early education: The constructivist view*. White Plains, NY: Longman.
- Diamond, A. (2010). The Evidence Base for Improving School Outcomes by Addressing the Whole Child and by Addressing Skills and Attitudes, Not Just Content. *Early Education and Development*, 21(5), 780-793. doi:10.1080/10409289.2010.514522
- Diamond, A., Barnett, W. S., Thomas, J., & Munro, S. (2007). Preschool Program Improves Cognitive Control. *Science*, 318(5855), 1387-1388. doi:10.1126/science.1151148
- Dotterer, A. M., Burchinal, M. R., Bryant, D., Early, D. M., & Pianta, R. (2009). Comparing universal and targeted prekindergarten programs. In R. Pianta & C. Howes (Eds.), *The promise of pre-k* (pp. 65-78). Baltimore, MD: Paul H. Brookes Publishing Co.
- Duncan, G. J., & Magnuson, K. (2013). Investing in Preschool Programs. *The Journal of Economic Perspectives*, 27(2), 109-132. doi:10.1257/jep.27.2.109
- Dunn, L. M., & Dunn, L. M. (1997). *Peabody Picture Vocabulary Test--Third Edition (PPVT-III)*. Upper Saddle River, NJ: Pearson Publishing.
- Early, D. M., Barbarin, O. A., Bryant, D., Burchinal, M. R., Chang, F., Clifford, R. M., . . . Kraft-Sayre, M. (2005). Pre-kindergarten in eleven states: NCEDL's multi-state study of pre-kindergarten and state-wide early educational programs (SWEEP) study. *National Institute for Early Education Research, Washington, DC*.

- Elkind, D. (2007). *The hurried child: Growing up too fast too soon*. Cambridge, MA: Da Capo Press.
- Fantuzzo, J. W., Gadsden, V. L., & McDermott, P. A. (2011). An Integrated Curriculum to Improve Mathematics, Language, and Literacy for Head Start Children. *American Educational Research Journal*, 48(3), 763-793. doi:10.3102/0002831210385446
- Goffin, S. G., & Wilson, C. (1994). *Curriculum models and early childhood education: Appraising the relationship*. New York, NY: Merrill
- Gormley, W. T., Phillips, D., & Gayer, T. (2008). Preschool Programs Can Boost School Readiness. *Science*, 320(5884), 1723-1724. doi:10.1126/science.1156019
- Harms, T., Clifford, R. M., & Cryer, D. (1998). *Early childhood environment rating scale*. New York: Teachers College Press.
- Herman, J. L., Klein, D. C. D., & Abedi, J. (2000). Assessing Students' Opportunity to Learn: Teacher and Student Perspectives. *Educational Measurement: Issues and Practice*, 19(4), 16-24. doi:10.1111/j.1745-3992.2000.tb00042.x
- Hightower, A. D. (1986). The Teacher–Child Rating Scale: A brief objective measure of elementary children's school problem behaviors and competencies. *School Psychology Review*.
- Hofferth, S. L. (1994). Who enrolls in head start? A demographic analysis of head start-eligible children. *Early Childhood Research Quarterly*, 9(3, Äi4), 243-268. doi:[http://dx.doi.org/10.1016/0885-2006\(94\)90009-4](http://dx.doi.org/10.1016/0885-2006(94)90009-4)
- Hulsey, L. K., Aikens, N., Kopack, A., West, J., Moiduddin, E., & Tarullo, L. B. (2011). *Head Start children, families, and programs: Present and past data from FACES*. (OPRE Report 2011-33a). Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services, Washington, DC.
- Jenkins, J. M. (2014). Early childhood development as economic development: Considerations for state-level policy innovation and experimentation. *Economic Development Quarterly*, 28(1), 147-165.
- Jenkins, J. M., Farkas, G., Duncan, G. J., Burchinal, M. R., & Vandell, D. (2016). Head Start at ages 3 and 4 versus Head Start followed by state pre-k: Which is more effective? *Educational Evaluation and Policy Analysis*, 38(1), 88-112.
- Karoly, L., Ghosh-Dastidar, B., Zellman, G. L., Perlman, M., & Fernyhough, L. (2008). *Prepared to learn: The nature and quality of early care and education for preschool-age children in California*. RAND, Santa Monica, CA.
- Morris, P. A., Mattera, S. K., Castells, N., Bangser, M., Bierman, K. L., & Raver, C. C. (2014). *Impact findings from the Head Start CARES Demonstration: National evaluation of three approaches to improving preschoolers' social and emotional competence.*, (OPRE Report 2014-44). Office of Planning Research and Evaluation, Administration for Children and Families., Washington, DC.
- National Center on Quality Teaching and Learning. (2014). *Preschool Curriculum Consumer Report*. Retrieved from <http://eclkc.ohs.acf.hhs.gov/hslc/tta-system/teaching/practice/curricula>
- Office of Head Start. (2010). Head Start Program Fact Sheet Fiscal Year 2010. Retrieved from <http://eclkc.ohs.acf.hhs.gov/hslc/data/factsheets/fHeadStartProgr.htm>
- Phillips, D. A., Gormley, W. T., & Lowenstein, A. E. (2009). Inside the pre-kindergarten door: Classroom climate and instructional time allocation in Tulsa's pre-K programs. *Early*

- Childhood Research Quarterly*, 24(3), 213-228.
doi:<http://dx.doi.org/10.1016/j.ecresq.2009.05.002>
- Piaget, J. (1976). *Piaget's theory*. New York, NY: Springer.
- Polikoff, M. S. (2015). How Well Aligned Are Textbooks to the Common Core Standards in Mathematics? *American Educational Research Journal*, 52(6), 1185-1211.
- Porter, A. C. (2002). Measuring the content of instruction: Uses in research and practice. *Educational Researcher*, 31(7), 3-14.
- Preschool Curriculum Evaluation Research Consortium. (2008). *Effects of preschool curriculum programs on school readiness*. (NCER 2008-2009). National Center for Education Research, Institute for Education Sciences, U.S. Department of Education, Washington, DC.
- Puma, M., Bell, S., Cook, R., Heid, C., Broene, P., Jenkins, F., . . . Downer, J. T. (2012). *Third grade follow-up to the Head Start Impact Study: Final Report*. (OPRE Report 2012-45). Washington, DC.
- Ritchie, S., Howes, C., Kraft-Sayre, M., & Weiser, B. (2001). *Emerging Academic Snapshot*. Unpublished measure, University of California at Los Angeles.
- Ritchie, S., & Willer, B. (2008). *Curriculum: A guide to the NAEYC early childhood program standard and related accreditation criteria*. Washington, D.C.: National Association for the Education of Young Children (NAEYC).
- Rosenbaum, P. R. (2002). *Observational studies*. New York: Springer-Verlag.
- Schmidt, W. H., Wang, H. C., & McKnight, C. C. (2005). Curriculum coherence: an examination of US mathematics and science content standards from an international perspective. *Journal of Curriculum Studies*, 37(5), 525-559. doi:10.1080/0022027042000294682
- Schweinhart, L. J. (2005). *Lifetime Effects: The High/Scope Perry Preschool Study through Age 40*. Ypsilanti, M.I.
- U.S. Department of Education. (2013). Early childhood education intervention report: The Creative Curriculum for Preschool, Fourth Edition. *Institute of Education Sciences, What Works Clearinghouse*.
- Vygotsky, L. S. (1978). Interaction between learning and development. *Readings on the development of children*, 23(3), 34-41.
- Walters, C. (2014). Inputs in the Production of Early Childhood Human Capital: Evidence from Head Start. *National Bureau of Economic Research Working Paper Series, No. 20639*. doi:10.3386/w20639
- Weikart, D. P., & Schweinhart, L. (1987). The High/Scope cognitively oriented curriculum in early education. In J. L. Roopnarine & J. E. Johnson (Eds.), *Approaches to early childhood education* (pp. 253-268). New York: Merrill/Macmillan.
- Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock-Johnson-III Tests of Achievement*. Itasca, IL: Riverside.
- Zigler, E. F., & Bishop-Josef, S. J. (2006). The cognitive child versus the whole child: Lessons from 40 years of Head Start. In D. G. Singer, R. Golinkoff, & K. Hirsh-Pasek (Eds.), *Play= learning: How play motivates and enhances children's cognitive and social-emotional growth* (pp. 15-35). New York, NY: Oxford University Press.
- Zill, N. (1990). *Behavior problems index based on parent report*. Washington, DC: Child Trends.

Table 1. Preschool Curricula Descriptions, Costs, and Dataset Availability

Curriculum	Publisher	Description	Cost Per Classroom	PCER	NCEDL	HSIS
Creative Curriculum	Teaching Strategies, Inc.	The Creative Curriculum® for Preschool focuses on project-based investigations as a means for children to apply skills. The curriculum is designed to foster development of the whole child through teacher-led, small and large group activities centered around 11 interest areas. The curriculum provides teachers with details on child development, classroom organization, teaching strategies, and engaging families in the learning process.	\$2149	X	X	X
HighScope	High/Scope Educational Research Foundation	HighScope® is based on the idea that children and adults learn best through hands-on experiences with people, materials, events, and ideas. Each individual program consists of a system of teaching practices, curriculum content areas for each topic and age group, assessment tools, and a training model.	\$1150	X	X	X
Scholastic Big Day for Pre-k	Scholastic	Big Day for PreK™ is organized into eight themes and includes five key elements of success: big experiences, meaningful conversations, best children's literature and nonfiction, innovative technology, and comprehensive program.	\$2900			X
DLM Early Childhood Express	McGraw-Hill	The DLM Early Childhood Express® is based on the idea that children learn best by connecting what they know with what they learn. The curriculum uses eight themed units to incorporate this philosophy into daily instruction.	\$4108	X		
HighReach Learning	Carson-Dellosa Publishing Group	HighReach Learning® is a project-based curriculum organized by topic. The curriculum that includes a 4-step planning process, which guides teachers in how to include children in the learning process. Additionally, family communication materials are included.	\$1125			X

Notes: Per-classroom estimates are approximated from the cost of purchasing the curriculum teacher's manual or equivalent (in 2015), and the baseline set of materials required to implement the curriculum. Publishers offer different sets of materials and thus costs will vary by publisher and curriculum.

Table 2. Comparison of Classroom Characteristics by Presence of Published Curricula in PCER and NCEDL

	PCER			NCEDL		
	Published Curriculum	No Published Curriculum	Diff	Published Curriculum	No Published Curriculum	Diff
Classroom Activities						
TBRS Math Quantity (0-3 scale)	1.22	.94	*	-	-	
TBRS Literacy Quantity (0-3 scale)	1.51	1.19	*	-	-	
Snapshot: Math Activity (proportion of day)	-	-		.06	.07	
Snapshot: Literacy Activity (proportion of day)	-	-		.15	.15	
Classroom Quality						
Arnett Caregiver Interaction	3.21	2.95	*	-	-	
Total ECERS Score	4.31	3.34	*	3.89	3.59	*
ECERS Factor 1 Language/Interactions	4.94	3.91	*	4.52	4.31	
ECERS Factor 2 Provisions for Learning	4.32	3.26	*	3.98	3.46	*
Teacher Characteristics						
Years of Teaching Experience	11.55	12.60		12.58	14.31	
Annual Salary/Hourly Wage	37595.68	38146.58		16.06	19.75	*
Female	.97	1.00		.98	.99	
Age	40.47	39.88		41.84	41.56	
No College/Less than a Bachelor's	.16	.25		.45	.33	
Associate's	.20	.10		-	-	
Bachelor's	.47	.40		.20	.15	
Master's/More than a Bachelor's	.16	.25		.34	.52	*
Center Characteristics						
Head Start	.34	.34		.35	.20	*
Public Pre-k	.64	.43	*	.65	.80	*
Private Child Care	.02	.24	*	-	-	
Full Day	-	-		.62	.22	*
Observations (Classrooms)	100	70		154	91	

Note. PCER = Preschool Curriculum Evaluation Research; NCEDL = National Center for Early Development and Learning; ECERS = Early Childhood Environment Rating Scale; TBRS = Teacher Behavior Rating Scale. Snapshot = Emerging Academics Snapshot. Hourly wage is measure of teacher compensation in NCEDL. Less than a Bachelor's is lowest level of education and More than a Bachelor's is highest level of education in NCEDL. In NCEDL, we only have enough data to distinguish between public pre-k and Head Start; there are no private child care centers that are not publicly-funded. * $p < .05$ from t-test for differences in means or Chi-squared test. All PCER classrooms observations rounded to the nearest 10 per NCES data security policy.

Table 3. Classroom Comparisons by Curricula in PCER, NCEDL, and HSIS

	HighScope	Creative Curriculum	DLM Express	No Published Curriculum	Diff	F-stat
PCER						
Classroom Activities (0-3 scale)						
TBRs Math Quantity	1.15	1.29	1.21	.94	*	4.06
TBRs Literacy Quantity	1.53	1.47	1.6	1.19	*	8.04
Classroom Quality						
Total ECERS Score	4.19	4.31	4.77	3.34	*	14.84
ECERS Factor 1 Language/Interactions	4.84	4.95	5.25	3.91	*	7.99
ECERS Factor 2 Provisions for Learning	4.33	4.22	4.65	3.26	*	16.34
Arnett Caregiver Interaction Score	3.12	3.3	3.25	2.95	*	4.19
Teacher characteristics						
Years of Teaching Experience	12.61	9.76	15.09	12.60		1.59
Annual Salary	44155.36	33412.11	31072.85	38146.58		1.89
No College	.10	.17	.36	.25		
Associate's	.15	.28	.09	.10		
Bachelor's	.66	.33	.36	.40	^	204.5
Master's	.10	.22	.18	.25		
Center characteristics						
Head Start	.17	.57	0	.34	*	6.53
Public Pre-k	.83	.39	1	.43	*	9.61
Private Child Care	0	.04	0	.24	*	5.47
Observations (Classrooms)	40 (23%)	50 (29%)	10 (6%)	70 (41%)		
	HighScope	Creative Curriculum	Other Published Curriculum	No Published Curriculum	Diff	F-stat
NCEDL						
Classroom Activities (in proportion of day)						
Snapshot: Math Activity	.05	.07	.08	.07	*	3.21
Snapshot: Literacy Activity	.15	.14	.16	.15		.37
Classroom Quality						
Total ECERS Score	3.95	3.76	3.91	3.60	*	2.92
ECERS Factor 1 Language/Interactions	4.60	4.18	4.70	4.33		1.71
ECERS Factor 2 Provisions for Learning	4.05	3.92	3.87	3.46	*	6.87
Teacher Characteristics						
Years of Teaching Experience	12.55	13.71	11.34	14.53		1.10
Hourly Wage	17.15	12.79	17.03	20.46	*	3.58
High School Diploma & Training Certificate	.06	.03	.06	0		1.78
Some College	.24	.26	.16	.08		
Associate's	.19	.18	.13	.20	^	27.96
Bachelor's	.18	.21	.25	.15		

DISTINCTIONS WITHOUT A DIFFERENCE

At Least 1 Year Past a Bachelor's	.11	.18	.16	.20			
Master's	.21	.13	.22	.29			
Center Characteristics							
Full Day	.66	.57	.59	.22	*	13.76	
Head Start	.39	.34	.25	.19	*	3.08	
Observations (Classrooms)	84 (36%)	38 (16%)	32 (14%)	80 (34%)			
	HighScope	Creative Curriculum	High Reach	Scholastic	Other Published Curriculum	Diff	F-stat
HSIS							
Classroom Activities (in times per month)							
Total Math Activities	107.88	114.33	119.50	125.52	106.65	*	6.36
Total Literacy Activities	151.40	153.74	167.33	160.11	152.62		1.28
Classroom Quality							
Total ECERS Score	5.34	5.04	4.69	4.92	5.01	*	7.79
Arnett Caregiver Interaction Score	2.56	2.49	2.67	2.51	2.54	*	2.83
Teacher characteristics							
Years of Teaching Experience	12.98	13.08	10.73	10.04	12.31		2.04
High School or Less	.02	.06	.03	.04	.04		
Some College	.01	.01	.03	0	.01		
Associate's	.28	.19	.15	.25	.20	^	44.22
Bachelor's	.37	.35	.65	.58	.36		
More than a Bachelor's	.32	.39	.15	.13	.39		
Center characteristics							
Full Day	.70	.68	.33	.64	.69	*	5.29
Community-Based Organization	.42	.39	.03	0	.47	*	13.84
Government Organization	.36	.41	.41	.85	.34	*	11.62
Public School	.17	.17	.03	.02	.14	*	2.86
Other Organization	.05	.03	.53	.12	.05	*	36.37
Observations (Classrooms)	334 (34%)	401 (40%)	38 (4%)	57 (6%)	167 (17%)		

Note. PCER = Preschool Curriculum Evaluation Research; ECERS = Early Childhood Environment Rating Scale; TBRS = Teacher Behavior Rating Scale; NCEDL = National Center for Early Development and Learning; Snapshot = Emerging Academics Snapshot; HSIS = Head Start Impact Study. * $p < .05$ from ANOVA. ^ $p < .05$ from Pearson Chi-Squared test. All PCER classroom observations rounded to the nearest 10 per NCES data security policy.

Table 4. NCEDL and HSIS Fixed Effects Results for Associations between Curricula Package and Children's School Readiness

	NCEDL				
	PPVT	WJAP	HT Competency	HT Problem Behaviors	
High Scope	0.06 (0.09)	-0.05 (0.14)	0.09 (0.13)	0.06 (0.13)	
Published Curriculum	0.15 (0.11)	-0.01 (0.15)	0.10 (0.15)	-0.15 (0.14)	
No Published Curriculum	0.18* (0.09)	0.05 (0.14)	0.00 (0.14)	0.23 ⁺ (0.14)	
Observations	398	394	452	450	
	HSIS				
	PPVT	WJAP	WJLW	WJSP	Behavior Problems
HighScope	0.01 (0.08)	0.20* (0.09)	0.08 (0.12)	0.30* (0.12)	0.02 (0.16)
High Reach	-0.24* (0.09)	0.09 (0.19)	0.36+ (0.20)	0.14 (0.21)	-0.40 (0.26)
Scholastic	0.15* (0.07)	0.30** (0.10)	0.25* (0.12)	0.52** (0.18)	0.14 (0.15)
Other Published Curriculum	0.04 (0.08)	0.15+ (0.09)	0.01 (0.10)	0.28+ (0.14)	0.25+ (0.15)
Observations	1592	1583	1594	1592	1556

Note. Clustered standard errors in parentheses. Creative Curriculum is reference group. All outcomes are in standard deviation units. Sample includes treated children from the HSIS experiment dataset only. All models include: child race, gender, age, and baseline assessment score for each outcome, mother's education, classroom quality (ECERS) and teacher controls (teacher's education and years of experience). Teen mom status was included in HSIS analyses (does not exist in NCEDL). An indicator for income under 150% of the poverty line was included in the NCEDL analyses (all HSIS participants were considered poor). PPVT = Peabody Picture Vocabulary Test, WJAP = Woodcock Johnson Applied Problems, WJLW = Woodcock Johnson Letter Word, WJSP = Woodcock Johnson Spelling, HT = Hightower. For both Problem Behaviors scores, a higher score indicates a more serious problem. + $p < .10$ * $p < .05$ ** $p < .01$

Figure 1. Study Conceptual Model

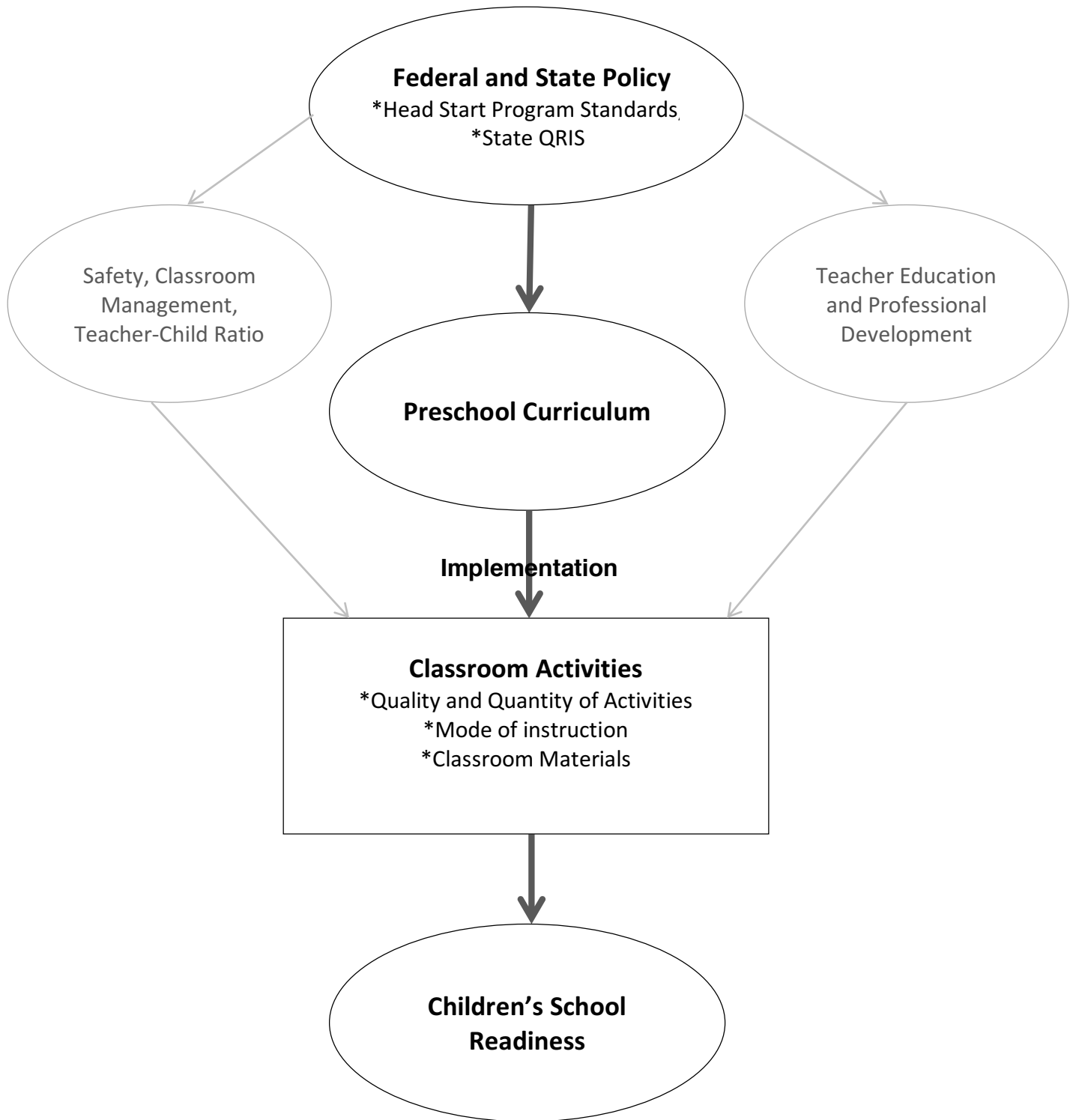
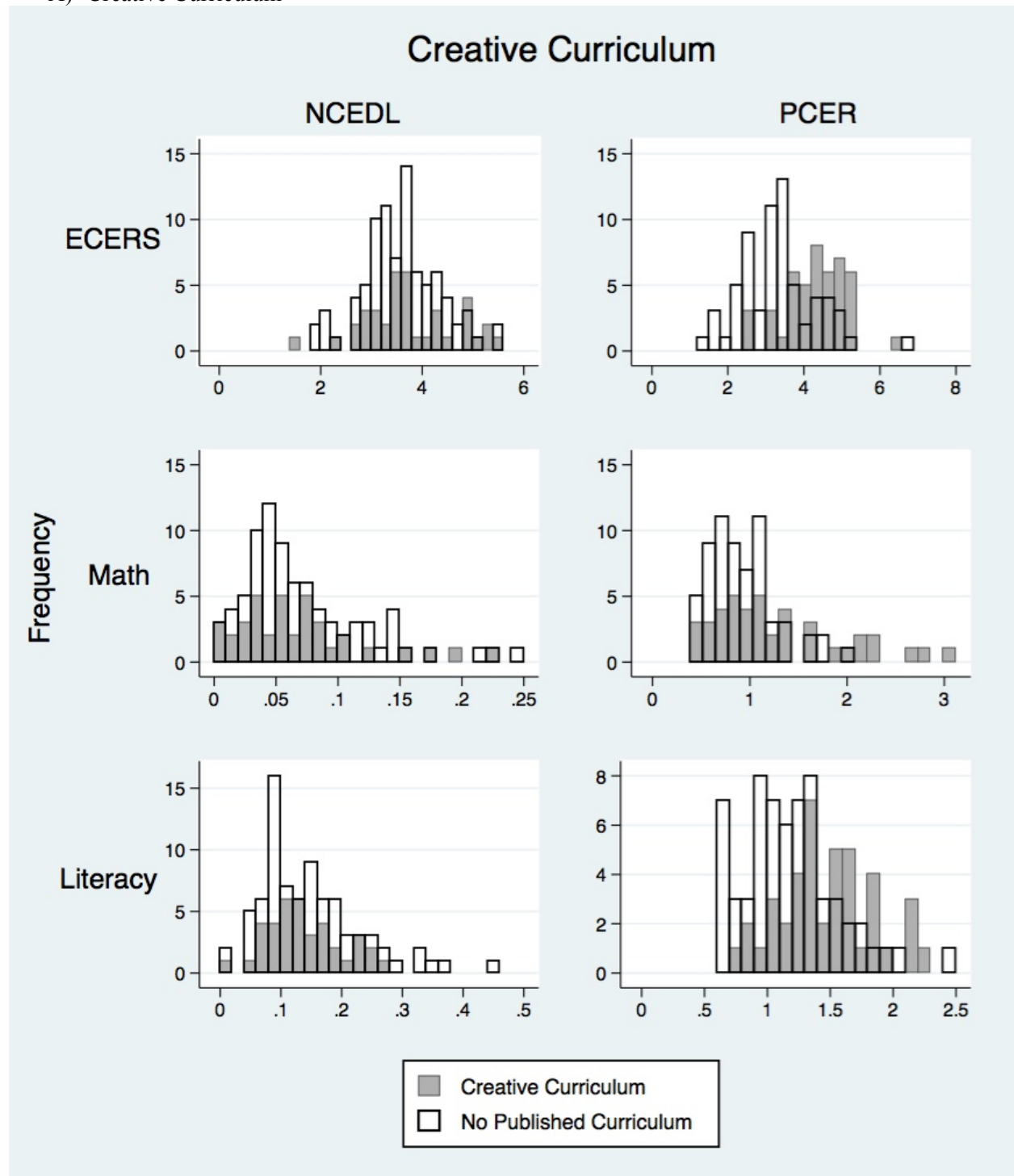
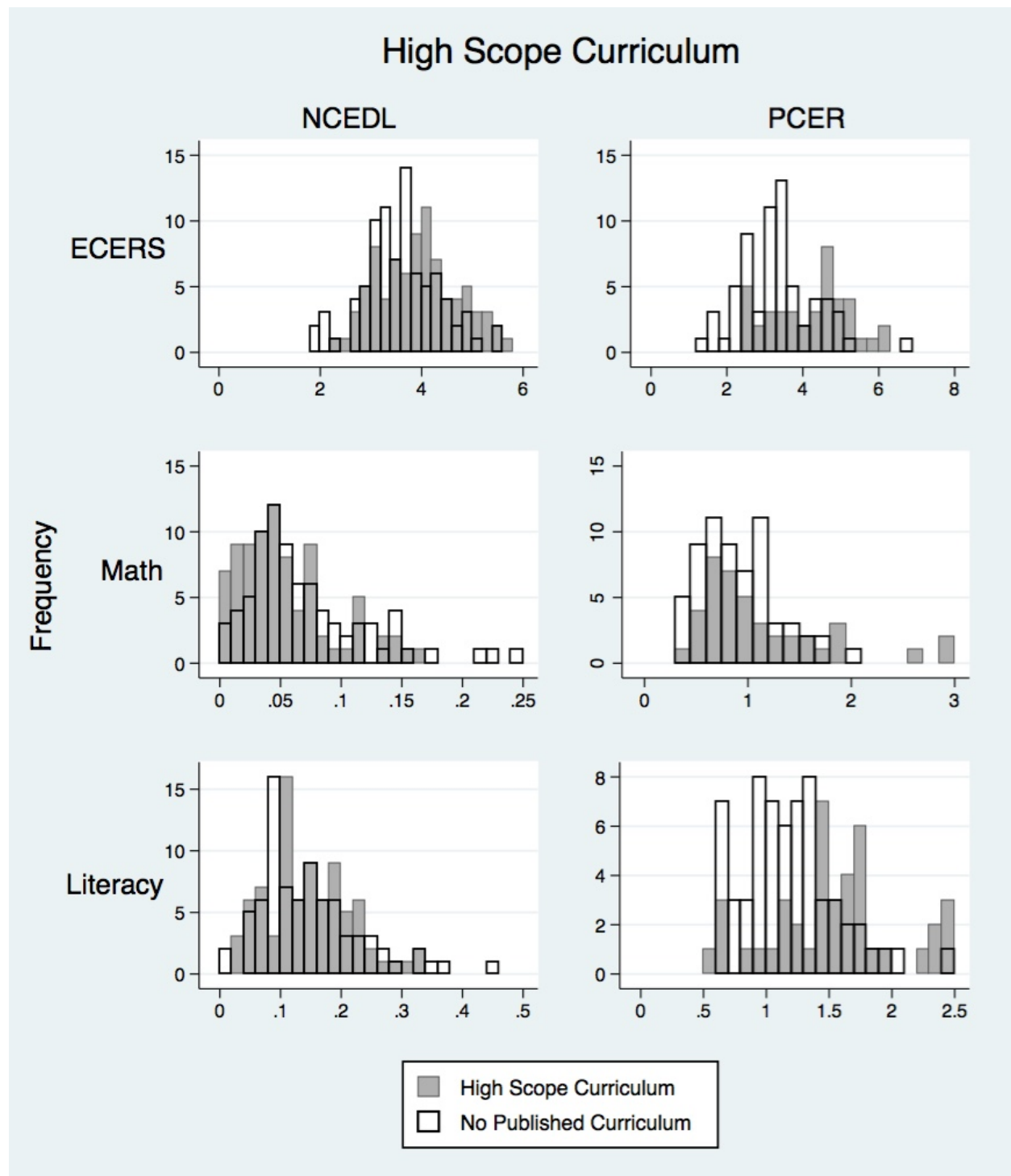


Figure 2: Histograms of Classroom Quality and Classroom activities in Creative Curriculum and High Scope Classrooms in the PCER and NCEDL Studies

A) Creative Curriculum



B) High Scope Curriculum



Note: Bins are comprised of classroom-level observations. ECERS, Math, and Literacy activities measures are in raw scale form (X-axis labels are omitted). ECERS scale ranges from 0-7, Math and literacy activities are shown as proportion of day in NCEDL (EAS Snapshot), and from a 0-3 scale in PCER (TBR).

Appendix A1. Classroom Characteristics by Curricula for PCER Sample

	High Scope	Creative Curriculum	DLM Express	No Published Curriculum	Diff	F-stat
Classroom Instruction (Proportion of classroom observations where 1=observed)						
Elicits responses during book reading	0.39	0.63	0.64	0.65	*	2.22
Visual rep. of vocabulary words	0.46	0.30	0.64	0.16	*	4.76
Asks open ended questions	0.68	0.70	0.73	0.47	*	2.55
Engage in letter knowledge	0.76	0.76	0.82	0.74		0.17
Express ideas in sentences	0.85	0.85	1.00	0.68	*	2.69
Scaffolding	0.88	0.89	1.00	0.62	*	6.34
Relates learned concepts/words	0.80	0.74	0.91	0.53	*	4.13
Talking with adults	0.98	0.98	1.00	0.84	*	2.95
Conversation w multiple turns	0.88	0.98	1.00	0.74	*	4.24
Vocabulary words singled out	0.56	0.37	0.82	0.29	*	4.34
Reads with expression	0.83	0.76	0.82	0.57	*	3.23
Flow and pacing of book reading	0.71	0.54	0.64	0.43	*	2.51
Extension activities from story	0.59	0.43	0.73	0.34	*	2.79
Compare letters	0.56	0.52	0.82	0.37	*	2.82
Discuss print	0.46	0.46	0.36	0.28		1.74
Writing opportunities	0.80	0.80	0.82	0.75		0.25
Rhyming	0.37	0.43	0.73	0.30	*	2.13
Alliteration	0.17	0.31	0.18	0.19		0.81
Sentence segmenting	0.05	0.02	0.09	0.02		0.56
Syllable blending/segmenting	0.20	0.17	0.27	0.03	*	2.51
Onset rime blending/segmenting	0.05	0.05	0.18	0.02		1.56
Phoneme blending/segmenting	0.17	0.17	0.27	0.03	*	2.33
Hands-on math	0.63	0.72	0.64	0.50		1.62
Incorporate math in daily routine	0.63	0.76	0.73	0.53	*	2.16
Math talk w. manipulatives	0.41	0.57	0.55	0.24	*	3.89
Teachable moments-math	0.59	0.70	0.73	0.37	*	4.18
Discussion of math concepts	0.44	0.57	0.27	0.21	*	4.69
Classroom Instruction (Frequency of activity on 1-4 scale)						
Work on learning the names of letters	1.32	1.33	1.00	1.35		0.64
Practice writing letters of the alphabet	1.76	2.11	2.18	1.76		0.95
Discuss new words	1.68	1.78	1.64	1.50		0.98
Dictate stories	2.51	2.84	2.45	2.35		1.33
Work on phonics	1.93	2.20	1.00	2.03		1.66
Listen to stories and show students print letters	1.32	1.36	1.45	1.30		0.16
Listen to stories without showing print	3.88	3.53	3.18	3.80		0.72
Retell stories	1.88	2.49	2.27	2.14		1.96
Learn about conventions of print	1.66	1.40	1.73	1.67		0.76
Students write their own names	1.32	1.67	1.82	1.29	*	2.29
Learn rhyming words	2.22	2.51	2.18	2.44		0.56
Learn about common prepositions	2.05	2.44	2.27	2.32		0.78
Count out loud	1.27	1.02	1.18	1.08	*	2.21
Work with geometric manipulatives	1.95	2.11	2.09	1.88		0.42
Work with counting manipulatives	2.05	1.93	1.73	1.77		0.83
Play math-related games	2.12	2.14	2.00	2.11		0.04
Use music to understand math ideas	2.68	2.44	2.55	2.42		0.28
Use creative movement to understand	2.98	2.91	3.18	2.79		0.32

Online Appendix

math ideas						
Work with rulers or measuring instruments	2.29	2.87	3.18	2.94	*	2.07
Engage in calendar-related activities	1.59	1.33	1.36	1.62		0.46
Classroom Quality (standardized)						
TBRS Math instructional quality	0.05	0.26	0.20	-0.40	*	3.72
TBRS Math instructional quantity	0.04	0.32	0.16	-0.36	*	3.30
TBRS Literacy instructional quality	0.25	-0.04	0.39	-0.61	*	7.39
TBRS Literacy instructional quantity	0.18	0.06	0.33	-0.57	*	6.14
Arnett Caregiver Interaction total	-0.01	0.33	0.24	-0.34	*	4.81
ECERS total score	0.24	0.36	0.79	-0.57	*	11.52
Teacher Characteristics						
Yrs. teaching experience	12.61	9.76	15.09	12.60		1.59
Annual salary	44155	33412	31072	38146		1.89
Female	0.95	0.98	1.00	1.00		0.92
Age	42.20	38.20	43.36	39.88		1.01
White	0.45	0.67	0.82	0.54	*	2.22
Black	0.33	0.26	0.18	0.34		0.97
Hispanic	0.13	0.02	0.00	0.03		1.77
Asian	0.00	0.00	0.00	0.01		0.35
Other race	0.10	0.04	0.00	0.07		0.50
No college	0.10	0.17	0.36	0.25		1.53
Associates	0.15	0.28	0.09	0.10		1.87
College	0.66	0.33	0.36	0.40	^	3.24
Masters	0.10	0.22	0.18	0.25		1.03
Center Characteristics						
Head Start	0.17	0.57	0.00	0.34	*	6.53
Public pre-k	0.83	0.39	1.00	0.43	*	9.61
Private child care	0.00	0.04	0.00	0.24	*	5.47
Observations (classrooms)	40	50	10	70		
Child and Family Characteristics						
White	0.32	0.24	0.24	0.36	*	5.48
Black	0.36	0.53	0.54	0.36	*	13.33
Hispanic	0.12	0.13	0.09	0.15		0.92
Male	0.47	0.56	0.52	0.51		2.28
<i>Preschool entry (baseline) assessments, standardized</i>						
PPVT	-0.03	-0.12	0.00	0.04		1.92
WJ Applied Problems	-0.04	-0.07	-0.20	0.06	*	2.68
WJ Letter word Identification	-0.23	0.07	0.00	-0.93	*	6.47
<i>Maternal Education</i>						
HS degree or equivalent	0.69	0.70	0.71	0.66		0.68
College degree or higher	0.08	0.09	0.22	0.16	*	7.82
Observations (Children)	380	360	100	600		

Note: * $p < .05$ from ANOVA. ^ $p < .05$ from Pearson Chi-Squared test. Baseline assessments are standardized to mean=0, standard deviation=1. All classroom observations rounded to the nearest 10 per NCES data security policy.

Appendix A2. Classroom Characteristics by Curricula for the NCEDL Sample

	High Scope	Creative Curriculum	Named Curriculum	No Published Curriculum	Diff	F-stat
Classroom Instruction (in proportion of day)						
Snapshot Activity Setting: Basics	0.20	0.20	0.19	0.19		0.73
Snapshot Activity Setting: Free Choice	0.34	0.36	0.39	0.33		1.16
Snapshot Activity Setting: Individ. Time	0.02	0.02	0.01	0.02		0.62
Snapshot Activity Setting: Meals/Snacks	0.14	0.14	0.11	0.13		1.25
Snapshot Activity Setting: Small Group Time	0.09	0.06	0.07	0.06		2.33
Snapshot Activity Setting: Whole Group Time	0.21	0.22	0.23	0.27	*	3.74
Snapshot Child Engagement: Aesthetics	0.18	0.19	0.18	0.16		1.32
Snapshot Child Engagement: Fine Motor	0.12	0.12	0.11	0.11		0.25
Snapshot Child Engagement: Gross Motor	0.08	0.08	0.09	0.07		1.47
Snapshot Child Engagement: Letters/Sounds	0.03	0.03	0.05	0.03		2.17
Snapshot Child Engagement: Math	0.05	0.07	0.08	0.07	*	3.21
Snapshot Child Engagement: Oral Lang Dev.	0.06	0.05	0.07	0.06		0.68
Snapshot Child Engagement: Read to	0.03	0.04	0.04	0.03		0.83
Snapshot Child Engagement: Pre Read	0.04	0.04	0.03	0.05		1.78
Snapshot Child Engagement: Science	0.09	0.08	0.12	0.11		2.43
Snapshot Child Engagement: Soc. Studies	0.19	0.18	0.20	0.18		0.19
Snapshot Child Engagement: Writing	0.01	0.01	0.01	0.01		2.38
Snapshot: Child distracted	0.03	0.04	0.02	0.04		1.93
Snapshot Adult Interaction: Routine	0.01	0.01	0.01	0.01		1.15
Snapshot Adult Interaction: Minimal	0.04	0.03	0.04	0.03		1.45
Snapshot Adult Interaction: Simple	0.07	0.06	0.06	0.06		0.53
Snapshot Adult Interaction: Elaborated	0.06	0.07	0.07	0.06		0.41
Snapshot Teacher-Child Engage: Encourages	0.19	0.17	0.16	0.17		0.90
Snapshot Teacher-Child Engage: Scaffolds	0.05	0.04	0.05	0.06		1.17
Snapshot Teacher-Child Engage: Didactic	0.14	0.15	0.18	0.18	*	4.14
Classroom Quality						
Total ECERS-R Score	3.95	3.76	3.91	3.60	*	2.92
ECERS-R Factor 1 Language/Interactions	4.60	4.18	4.70	4.33		1.71
ECERS-R Factor 2 Provisions for Learning	4.05	3.92	3.87	3.46	*	6.87
CLASS: Positive Climate	5.05	4.83	5.19	5.14		1.21
CLASS: Negative Climate	1.82	1.93	1.65	1.54	*	2.92
CLASS: Teacher Sensitivity	4.52	4.14	4.61	4.55		1.66
CLASS: Overcontrol	1.90	1.87	1.88	1.93		0.04
CLASS: Behavioral Management	4.82	4.52	4.95	4.85		1.39
CLASS: Productivity	4.40	4.20	4.61	4.43		1.02
CLASS: Concept Development	2.14	1.95	2.46	2.18		1.47
CLASS: Learning Formats	4.27	3.76	4.35	4.11	*	3.10
CLASS: Roteness/Basic Skills	2.16	2.00	2.49	2.04		1.76
CLASS: Quality of Feedback Mean	1.65	1.64	1.75	1.81		0.88
CLASS: Total - Emot. & Instruc. Climate	4.35	4.14	4.49	4.40		1.87
CLASS: Factor 1 - Emotional Climate	5.34	5.14	5.45	5.42		1.42
CLASS: Factor 2 - Instructional Climate	1.89	1.80	2.10	1.99		1.03
Teacher Characteristics						
Years experience with kids prior to kindergarten	9.26	10.24	8.4	10.5		0.84
Total years experience	12.55	13.71	11.34	14.53		1.10
HS diploma or equivalent	0.00	0.00	0.00	0.01		0.64
HS diploma & training certificate	0.06	0.03	0.06	0.00	^	1.78
Some college	0.24	0.26	0.16	0.08	^	3.41

Online Appendix

Associates	0.19	0.18	0.13	0.20		0.30
College	0.18	0.21	0.25	0.15	^	0.57
At least 1 year past BA	0.11	0.18	0.16	0.20	^	0.96
Master's degree	0.21	0.13	0.22	0.29	^	1.24
Education specialist or professional degree	0.01	0.00	0.03	0.08	^	2.24
Child Development Associate	0.63	0.53	0.72	0.79	*	3.24
Female	0.99	0.95	1.00	0.99		1.21
Calculated teacher hourly wage	17.15	12.79	17.03	20.46	*	3.58
State certification to teach 4 year olds	0.67	0.77	0.72	0.74		0.58
Age	41.91	42.11	41.38	41.30		0.07
Classroom and Center Characteristics						
Class size in observed class - Spring	19.64	20.28	19	18.47		0.69
<i>Classroom Type</i>						
3 & 4 yr. olds	0.39	0.41	0.45	0.41		0.11
4 yr. olds (+ 5 yr. olds who will be in K next yr.)	0.58	0.59	0.55	0.59		0.06
4 year olds & kindergartners	0.02	0.00	0.00	0.00		1.18
Length of school day (hours)	5.76	5.71	5.83	3.81	*	12.41
Class is full day	0.66	0.57	0.59	0.22	*	13.76
Located in a public school	0.46	0.39	0.50	0.73	*	5.77
Head Start	0.39	0.34	0.25	0.19	*	3.08
Observations (Classrooms)	84	38	32	80		
Child and Family characteristics						
White	0.41	0.45	0.47	0.50		1.58
Black	0.33	0.25	0.35	0.22	*	4.94
Hispanic	0.29	0.30	0.22	0.32		1.56
Male	0.47	0.52	0.46	0.50		0.56
<i>Preschool entry (baseline) assessments, standardized</i>						
PPVT	-0.08	0.05	0.02	0.05		0.99
WJ Applied Problems	-0.12	0.03	0.03	0.010		2.30
Hightower Competency	0.00	0.05	0.06	-0.05		0.56
Hightower Problem Behavior	0.07	-0.10	-0.03	-0.01		1.03
<i>Maternal Education</i>						
Below high school	0.18	0.19	0.15	0.18		0.24
High school	0.22	0.21	0.19	0.25		0.91
Some college	0.48	0.45	0.44	0.42		0.74
College degree or higher	0.12	0.15	0.22	0.14		2.27
Observations (Children)	359	161	135	360		

Note: * p<.05 from ANOVA. ^p<.05 from Pearson Chi-Squared test. Baseline assessments are standardized to mean=0, standard deviation=1.

Appendix A3. Characteristics of Classrooms by Curricula in the HSIS Sample

	High Scope	Creative Curriculum	High Reach	Scholastic	Other Curriculum	Diff	F- stat
Classroom Instruction (times per month)							
Total literacy activities	151.4	153.74	167.33	160.11	152.62		1.28
Total math activities	107.88	114.33	119.5	125.52	106.65	*	6.36
Listen to stories (no print)	8.45	6.50	5.61	5.16	7.41	*	3.67
Retell or makeup stories	10.60	10.76	13.15	11.96	9.98		2.28
Write letters of alphabet	11.95	12.62	12.94	13.04	12.83		0.78
Learn names of letters	15.85	15.93	17.04	16.96	16.11		0.75
Discuss new words	13.70	15.17	16.17	15.55	14.84	*	3.48
Have children tell you a story	10.61	11.66	13.85	13.63	10.39	*	5.03
Practice sounds letters make	13.65	13.38	17.42	11.77	13.92	*	3.96
Listen to stories w. print	15.93	16.79	18.23	18.49	16.15	*	3.77
Show child how to read book	15.09	15.01	16.09	17.20	15.44		1.64
Write own name	14.37	14.41	14.86	14.29	15.78		1.70
Rhyming words and families	9.96	9.88	15.07	11.21	9.93	*	5.42
Teach directional words like over and up	12.79	13.54	14.15	13.30	12.15		1.65
Count out loud	18.22	18.84	19.68	19.23	17.87	*	4.24
Work with shape blocks	16.06	17.26	18.86	17.44	15.06	*	7.80
Count small toys to learn math	15.22	16.71	16.49	17.02	15.87	*	4.31
Play math games	11.41	13.28	14.01	15.79	13.26	*	7.89
Use music to understand math ideas	10.73	11.39	13.09	13.24	10.18	*	3.26
Use dance to practice math ideas	9.66	10.17	12.64	13.31	9.51	*	5.15
Work with rulers/measuring cups	10.62	10.25	8.90	10.16	9.32		1.17
Talk about calendar/days of week	17.18	17.34	19.68	19.89	16.95	*	4.52
Arts and crafts activities	17.26	17.41	16.03	18.84	17.35	*	4.74
Play with games/toys indoors	19.10	19.27	18.95	19.54	18.69	*	6.58
Play sports/exercise	17.91	17.87	18.33	19.22	17.77	*	3.59
Child chooses activities (hrs./day)	2.22	2.18	1.96	1.74	1.97		2.00
Adult directs individual. activities (hrs./day)	0.76	0.98	1.24	0.96	1.16		1.40
Adult directs small group (hrs./day)	0.64	0.86	1.14	0.83	0.97		1.39
Adult directs whole class (hrs./day)	0.91	0.95	1.44	1.13	1.19		1.88
Classroom Quality							
Arnett Caregiver Interaction total score	76.47	74.49	80.05	75.15	75.95	*	2.89
ECERS-R activity mean score	4.81	4.54	3.76	4.37	4.38	*	11.40
ECERS-R activities and program mean score	5.07	4.78	3.81	4.58	4.65	*	13.60
ECERS-R interaction mean score	5.86	5.55	5.72	5.42	5.68	*	3.57
ECERS-R language reasoning mean score	5.38	5.01	5.05	5.00	5.01	*	4.56
ECERS-R personal care routines mean score	5.58	5.30	5.19	5.14	5.30	*	2.64
ECERS-R program structure mean score	5.84	5.52	4.41	5.25	5.37	*	9.38
ECERS-R space and furnishings mean score	5.27	5.01	4.79	4.99	5.05	*	4.17
ECERS-R Total score	5.34	5.04	4.69	4.92	5.01	*	7.79
Teacher characteristics							
Yrs. teaching Head Start	8.00	9.22	7.39	6.05	7.17	*	3.65
Yrs. teaching experience	12.98	13.08	10.73	10.04	12.31		2.04
Child Development Associates	0.59	0.56	0.64	0.70	0.55		1.37
HS or below	0.02	0.06	0.03	0.04	0.04		
Some college	0.01	0.01	0.03	0.00	0.01		
Associates	0.28	0.19	0.15	0.25	0.20	^	44.22
College	0.37	0.35	0.65	0.58	0.36		
College+	0.32	0.39	0.15	0.13	0.39		

Online Appendix

Classroom and Center characteristics							
Training in curriculum	0.96	0.92	1.00	0.96	0.95		2.00
Full day classroom	0.70	0.68	0.33	0.64	0.69	*	5.29
Community-based organization	0.42	0.39	0.03	0.00	0.47	*	13.84
Government organization	0.36	0.41	0.41	0.85	0.34	*	11.62
Public school	0.17	0.17	0.03	0.02	0.14	*	2.86
Other organization	0.05	0.03	0.53	0.12	0.05	*	36.37
Observations (Classrooms)	334	401	38	57	167		
Child and Family characteristics							
White	0.18	0.36	0.30	0.04	0.35	*	13.22
Black	0.72	0.22	0.38	0.10	0.30	*	30.54
Hispanic	0.10	0.42	0.33	0.87	0.35	*	37.46
Male	0.46	0.49	0.48	0.60	0.49		0.60
<i>Preschool entry (baseline) assessments, standardized</i>							
PPVT	-0.01	-0.02	0.09	-0.51	0.00	*	6.94
WJ Applied Problems	0.07	0.04	-0.09	-0.18	-0.02		1.87
WJ Letter word Identification	0.06	0.04	-0.15	-0.02	0.00		0.84
<i>Maternal Education</i>							
Below High School	0.4	0.39	0.35	0.57	0.34	*	5.36
HS degree or equivalent	0.26	0.32	0.35	0.21	0.38	*	3.34
Beyond High School	0.34	0.29	0.3	0.22	0.28		0.92
Teenage mom	0.25	0.16	0.15	0.18	0.16		1.44
Observations (Children)	626	803	68	104	282		

Note. * p<.05 from ANOVA. ^p<.05 from Pearson Chi-Squared test. Sample includes HSIS treatment children only, from both age-3 and age-4 cohorts. Baseline assessments are standardized to mean=0, standard deviation=1.

Appendix B. Classroom Outcomes by Curricular Status in PCER and NCEDL

	PCER					
	ECERS		Math Activities		Literacy Activities	
	(1)	(2) + State FE	(3)	(4) + State FE	(5)	(6) + State FE
Has Curriculum	0.32 (0.59)	0.01 (0.54)	-0.56 (0.49)	0.32 (0.59)	0.01 (0.54)	-0.56 (0.49)
Teacher's Total Years Experience	0.02 (0.01)	-0.00 (0.01)	-0.01 (0.01)	0.02 (0.01)	-0.00 (0.01)	-0.01 (0.01)
Teacher's Education: BA	0.30 (0.21)	0.09 (0.25)	0.04 (0.25)	0.30 (0.21)	0.09 (0.25)	0.04 (0.25)
Teacher's Education: More Than BA	0.40 (0.26)	0.39 (0.29)	0.29 (0.29)	0.40 (0.26)	0.39 (0.29)	0.29 (0.29)
Class Proportion Female	-1.05* (0.45)	-1.08* (0.53)	-1.33** (0.49)	-1.05* (0.45)	-1.08* (0.53)	-1.33** (0.49)
Class Proportion Latino	0.32 (0.76)	-0.31 (0.70)	-0.24 (0.67)	0.32 (0.76)	-0.31 (0.70)	-0.24 (0.67)
Class Proportion Black	-0.27 (0.37)	-0.18 (0.31)	-0.45 (0.32)	-0.27 (0.37)	-0.18 (0.31)	-0.45 (0.32)
Class Proportion Asian	-0.71 (0.92)	0.85 (1.40)	-1.78 (1.45)	-0.71 (0.92)	0.85 (1.40)	-1.78 (1.45)
Class Proportion Other Race	-0.98 (0.67)	-0.79 (0.87)	-1.71* (0.81)	-0.98 (0.67)	-0.79 (0.87)	-1.71* (0.81)
Class Mother Ed., College	-0.34 (0.45)	0.01 (0.45)	-0.81* (0.38)	-0.34 (0.45)	0.01 (0.45)	-0.81* (0.38)
Constant	-0.07 (0.60)	0.30 (0.60)	1.32* (0.60)	-0.07 (0.60)	0.30 (0.60)	1.32* (0.60)
Observations	130	130	130	130	130	130

	NCEDL					
	ECERS		Math Activities		Literacy Activities	
	(1)	(2) + State FE	(3)	(4) + State FE	(5)	(6) + State FE
Has Curriculum	0.53** (0.20)	0.39* (0.19)	-0.01 (0.19)	0.11 (0.20)	0.03 (0.22)	0.01 (0.25)
Teacher's Total Years Experience	0.01 (0.01)	0.01 (0.01)	0.03* (0.01)	0.03* (0.01)	0.01 (0.01)	0.01 (0.01)
Teacher's Education: BA	0.25 (0.51)	0.36 (0.60)	-0.07 (0.31)	0.27 (0.40)	-0.42 (0.42)	-0.35 (0.49)
Teacher's Education: More Than BA	0.28 (0.27)	0.58 (0.37)	-0.03 (0.25)	0.26 (0.32)	-0.20 (0.28)	-0.19 (0.32)
Class Proportion Female	0.07 (0.44)	0.10 (0.41)	0.28 (0.48)	-0.02 (0.43)	0.32 (0.45)	0.25 (0.48)
Class Proportion Latino	-0.58* (0.23)	-0.27 (0.29)	0.25 (0.22)	-0.07 (0.40)	0.07 (0.24)	-0.32 (0.38)
Class Proportion Black	-0.42 (0.31)	-0.28 (0.31)	0.45 (0.28)	0.55+ (0.28)	0.10 (0.26)	0.14 (0.27)
Class Proportion Other Race	0.04 (0.42)	0.19 (0.46)	1.19* (0.50)	1.07+ (0.61)	1.10* (0.45)	0.80 (0.53)
Class Mother Ed., College	0.11 (0.38)	0.03 (0.41)	0.97* (0.39)	1.11** (0.41)	0.29 (0.36)	0.37 (0.42)
Constant	-0.48 (0.38)	-0.83* (0.41)	-0.94* (0.46)	-1.13* (0.46)	-0.26 (0.35)	-0.11 (0.42)
Observations	143	143	139	139	139	139

Note. Standard errors in parentheses. Outcomes are standardized. Coefficients are relative to classrooms with no published curriculum. All models control for classroom proportion: female, Latino, black, other race, and mothers with a college degree. Teacher controls include total years of experience and highest level of education. ECERS = Early Childhood Environment Rating Scale. Math and literacy activities were measured with the Teacher Behavior Rating Scale for PCER and the Emerging Academics Snapshot for NCEDL. + $p < .10$, * $p < .05$, ** $p < .01$.

Appendix C. NCEDL State Fixed Effects Results of Child School Readiness by Curriculum Status

	PPVT		WJAP		HT Competency		HT Problem Behaviors	
	(1) State FE + Child Covariates	(2) + Teacher Covariates & Classroom Quality	(3) State FE + Child Covariates	(4) + Teacher Covariates & Classroom Quality	(5) State FE + Child Covariates	(6) + Teacher Covariates & Classroom Quality	(7) State FE + Child Covariates	(8) + Teacher Covariates & Classroom Quality
Has Curriculum	-0.01 (0.05)	-0.07 (0.06)	0.02 (0.07)	-0.05 (0.09)	0.04 (0.08)	0.07 (0.10)	-0.15* (0.07)	-0.24* (0.11)
Observations	733	419	719	415	837	480	835	478

Note. Standard errors are clustered at the classroom level. Outcomes are standardized. Classrooms with no curriculum are the reference group. All models include: state fixed effects child race, child gender, and whether the mother has a college degree. Teacher covariates include highest level of education and total years of experience. Classroom quality is measured by the spring ECERS score. PPVT = Peabody Picture Vocabulary Test, WJAP = Woodcock Johnson Applied Problems, HT = Hightower. For HT Problem Behaviors, a higher score indicates a more serious problem, ECERS = Early Childhood Environment Rating Scale. + $p < .10$ * $p < .05$ ** $p < .01$

Appendix D: Associations between Curricular Packages and Classroom-level Outcomes

PCER			
	(1) ECERS	(2) TBRS Math Activities	(3) TBRS Literacy Activities
High Scope	-0.32 (0.35)	-0.27 (0.38)	0.07 (0.23)
DLM Express	0.60** (0.19)	-0.58 (0.62)	0.08 (0.31)
No curriculum	-1.01* (0.37)	-0.58+ (0.32)	-0.61* (0.23)
Observations	160	150	150

Note. All outcome scales are in standard deviation units (mean 0, sd 1). Standard errors clustered at the site level (in parentheses). Creative Curriculum is the reference group. All models include: proportion of classroom female, Hispanic, black, Asian, or other race, mother's education, and teacher characteristics (hourly wage, education, experience). Classroom treatment status was also included as a covariate. Classroom N rounded to the nearest 10 per NCES data security policy. + $p < .10$ * $p < .05$ ** $p < .01$

NCEDL			
	(1) ECERS	(2) Snapshot Math	(3) Snapshot Literacy
High Scope	0.18 (0.23)	-0.47+ (0.27)	0.14 (0.22)
Other published curriculum	0.35 (0.25)	-0.03 (0.33)	0.50+ (0.28)
No published curriculum	-0.21 (0.25)	-0.34 (0.28)	0.18 (0.30)
Observations	143	139	139

Note. All outcome scales are in standard deviation units (mean 0, sd 1). Standard errors clustered at the state level in parentheses. Creative Curriculum is the reference group. All models include: proportion of classroom female, Hispanic, black, or other race, mother's education, and teacher characteristics (hourly wage, education, experience). + $p < .10$ * $p < .05$ ** $p < .01$

HSIS				
	(1) ECERS	(2) Total Math Activities	(3) Total Literacy Activities	(4) Arnett Caregiver Interaction Scale
HighScope	0.18 (0.16)	-0.17* (0.08)	-0.02 (0.11)	0.09 (0.17)
High Reach	-0.23 (0.24)	0.02 (0.12)	0.11 (0.16)	0.74** (0.25)
Scholastic	-0.40** (0.13)	0.32** (0.10)	0.20+ (0.10)	-0.17 (0.24)
Other curriculum	-0.09 (0.18)	-0.18 (0.14)	0.00 (0.12)	0.09 (0.17)
Observations	885	953	953	878

Note. All outcome scales are in standard deviation units (mean 0, sd 1). Standard errors clustered at the program level (in parentheses). Creative Curriculum is the reference group. All models include: proportion of classroom female, Hispanic, black, Asian, or other race, mother's education, and teacher characteristics (education, experience). + $p < .10$ * $p < .05$ ** $p < .01$

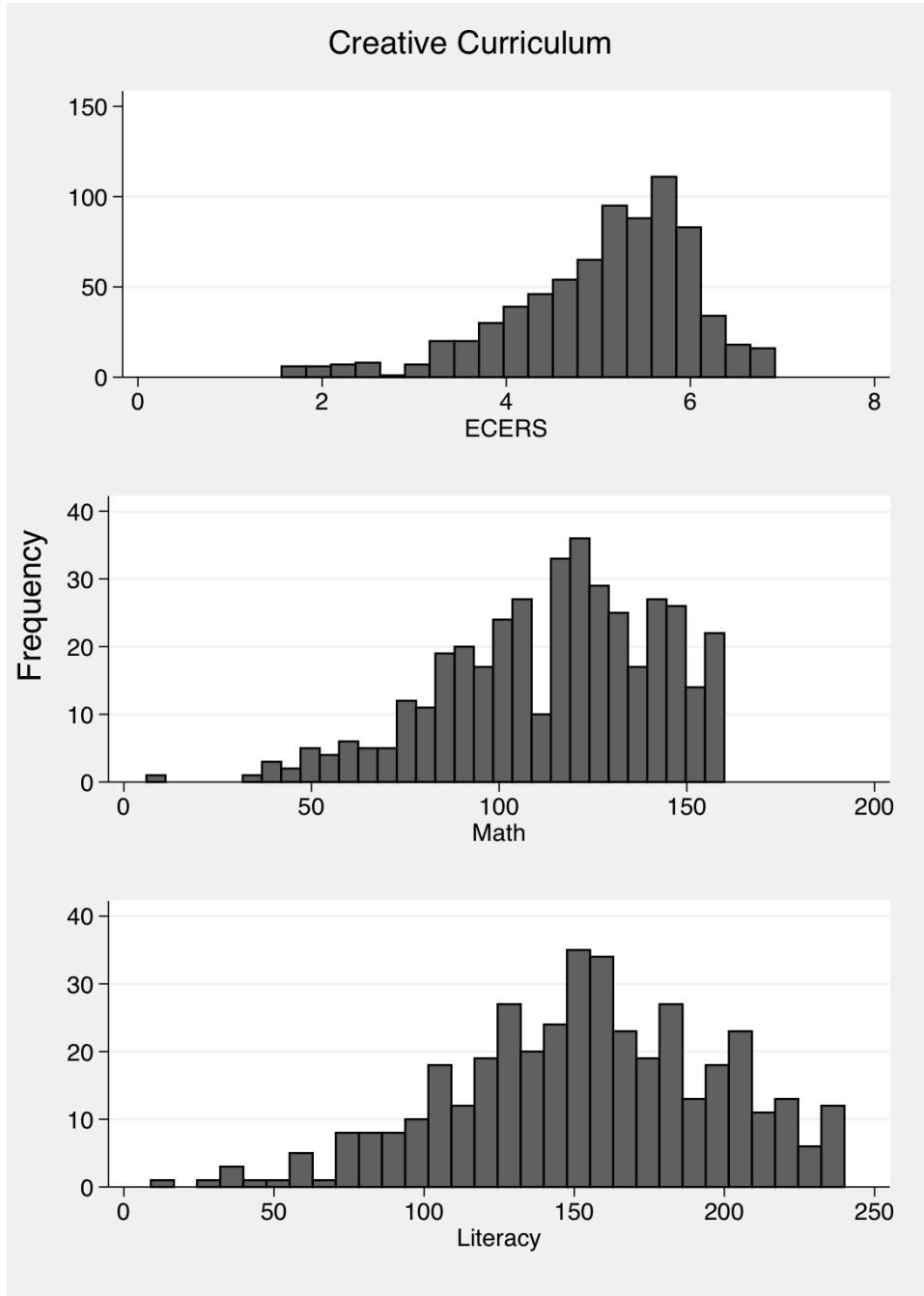
Appendix E. NCEdL State Fixed Effects Results for Curricular Package Associations with Child Outcomes with Additional Controls

	(1) PPVT	(2) WJAP	(3) HT Competency	(4) HT Problem Behaviors
High Scope	0.11 (0.10)	-0.06 (0.15)	0.15 (0.15)	0.11 (0.13)
Other Published Curriculum	0.18 (0.11)	-0.03 (0.15)	0.14 -0.17	-0.06 (0.15)
No Published Curriculum	0.19* (0.09)	-0.01 (0.15)	0.06 -0.18	0.36* (0.15)
Observations	341	337	386	384

Note. Standard errors clustered at the site level (in parentheses). Creative Curriculum is the reference group. All models include: state fixed effects, child and family characteristics (child race, child gender, child age, mother's education, poverty), classroom quality (ECERS), whether program was full day, whether program was Head Start, and teacher characteristics (hourly wage, education, experience). + p<0.10, *p<0.05, **p<0.01,

Appendix F: Histograms of Classroom Quality and Activities by Curriculum in the HSIS sample.

A) Creative Curriculum



B) High Scope Curriculum

