Published in final edited form as:

J Appl Dev Psychol. 2016 ; 43: 29–42. doi:10.1016/j.appdev.2015.12.007.

Differential Effectiveness of Head Start in Urban and Rural Communities

Dana Charles McCoy¹, Pamela A. Morris², Maia C. Connors², Celia J. Gomez¹, and Hirokazu Yoshikawa²

¹Harvard Graduate School of Education, Cambridge, MA 02138

²New York University, New York, NY 10003

Abstract

Recent research suggests that Head Start may be differentially effective in improving low-income children's early language and literacy skills based on a number of individual- and family-level characteristics. Using data from the Head Start Impact Study ($n = 3503$; 50% male, 63% treatment group), the present study extends this work to consider program impact variation based on centers' location in urban versus rural communities. Results indicate that Head Start is more effective in increasing children's receptive vocabulary (as measured by the PPVT) in urban areas and their oral comprehension (as measured by the Woodcock-Johnson Oral Comprehension task) in rural areas. Additional analyses suggest that related characteristics of the center – including concentration of dual language learners and provision of transportation services – may underlie these associations. Implications for research on program evaluation and policy are discussed.

Keywords

Head Start; preschool; impact variation; language and literacy; urbanicity; neighborhoods

Since the 1960's, increased understanding of the importance of the early childhood developmental period has led to substantial investment in early childhood care and education (ECCE) programs like Head Start for promoting school readiness and reducing income-based inequities at school entry. Rigorous evaluations of Head Start have shown mixed evidence for the program's effectiveness in achieving these goals, with some studies showing positive impacts on children's pre-academic skills (e.g., Deming, 2009; McKey, 1985; Shager et al., 2013) and others suggesting more modest or null effects (e.g., Bernary, 2012; Currie & Thomas, 1993). Although useful for quantifying overall effectiveness, these studies of average program impact are less helpful for identifying specific conditions under which Head Start may be particularly beneficial – or deleterious – for children (Bloom & Weiland, 2015). Given the heterogeneity of community settings in which Head Start serves

Address correspondence to: Dana Charles McCoy, Harvard Graduate School of Education, Larsen Hall, Room 704, 14 Appian Way, Cambridge, MA 02138, dana_mccoy@gse.harvard.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

children in the United States, understanding specific contextual sources of program impact variation is of critical importance for identifying existing programmatic strengths and weaknesses, providing more targeted approaches to addressing children's needs, and (re)allocating resources to optimize equity.

In the present study, we use data from the Head Start Impact Study (HSIS) to provide new, hypothesis-generating evidence on Head Start impact variation across urban and rural communities in the United States. First, we provide a descriptive characterization of the community settings in which HSIS Head Start centers are located. In particular, we focus on Head Start communities' levels of urbanicity, as defined by the percent of families within the surrounding census tract neighborhood who are living in urbanized areas or clusters with more than 400 people per square mile. Second, we explore whether the effectiveness of Head Start for promoting children's short-term early language and literacy outcomes – i.e., receptive vocabulary, oral comprehension, early writing, and early reading skills at the end of the preschool year – differs based on centers' levels of community urbanicity. Finally, we test whether any observed urban-rural differences in effectiveness may actually be explained by other contextual characteristics that are related to urbanicity. Specifically, we test whether larger impacts in urban or rural environments may be driven by co-varying levels of 1) neighborhood demographics, crime, and resource availability, 2) center characteristics and capacities, including teachers' average levels of education, centers' provision of family services, and quality of teacher-student interactions and 3) center compositional characteristics, including the types of families and children served.

Neighborhood Urbanicity and Head Start – Evidence for Contextual Strengths and Challenges

An extensive body of research from the psychological, sociological, and economic literatures suggests that neighborhoods play an important role in shaping children's early development (Aikens & Barbarin, 2008; Brooks-Gunn, Duncan, & Aber, 1997; Caughy, Hayslett-McCall, & O'Campo, 2007; Leventhal & Brooks-Gunn, 2000). Neighborhood socioeconomic disadvantage has been strongly linked with negative outcomes throughout the life trajectory through its direct and indirect effects on children's stress responses, family interactions, and broader social relationships (Boardman, Finch, Ellison, Williams, & Jackson, 2001; Chetty, Hendren, Kline, & Saez, 2014; Pinderhughes, Nix, Foster, Jones, & The Conduct Problems Prevention Research Group, 2001; Sampson, Morenoff, & Earls, 1999). Importantly, the ways that community poverty manifests across diverse geographical and sociocultural contexts appears to have important implications for its effects on children and families. The urban poor, for example, tend to report better physical health outcomes but worse psychological functioning than the rural poor, with these relationships differing for individuals from diverse sociodemographic and racial/ethnic backgrounds (Amato & Zuo, 1992; Bendar, Fedor, & Carlson, 2011; Glaeser, 2011; Rutter, 1980). Although the mechanisms underlying urban versus rural poverty's effects on children are relatively poorly understood, researchers have posited that differential concentrations of particular risk and protective factors like crime and social/educational resources across the urbanicity continuum may be responsible (Amato & Zuo, 1992; Rutter, 1980).

In this paper, we extend work on neighborhood settings to consider the ways that urban and rural environments – and the risk and protective factors therein – explain variation in the impact of Head Start on young children’s language and literacy development. Launched in 1965 across 2400 communities through President Lynden B. Johnson’s “War on Poverty,” Head Start has been conceptualized as a critical social resource for buffering children and families from the effects of both urban and rural poverty. Since its inception, Head Start has provided comprehensive educational, social, and health-related services to more than 30 million children across all 50 states, as well as the District of Columbia, Puerto Rico, and the U.S. territories (Office of Head Start, 2014). Although Head Start’s mission has always been to serve children across a diversity of communities – including urban, suburban, and rural environments facing various risk and protective factors – research on Head Start and related educational programs for low-income children has historically focused almost exclusively on low-income, urban environments (Tieken, 2014). The under-representation of rural communities in Head Start research has led to a dearth of knowledge regarding whether Head Start’s programmatic model is able to support children equally across different types of communities. Given that up to 30 percent of Head Start children are served in rural areas (Rural Poverty Research Institute, 2008), understanding the extent of urban-rural disparities is of great relevance to both advocacy and policy, as it would not only draw attention to existing inequities, but would also provide information on where additional resources are needed.

Descriptive research suggests several important differences across rural and urban environments that may affect Head Start programs’ ability to provide optimal services to children. Rural ECCE programs, for example, have traditionally been shown to be more personal and less bureaucratic than urban programs, yet face more difficulty recruiting highly credentialed staff and achieving “economies of scale” due to lower population density, difficulty in transportation, and reduced resources (Chertow, 1968; National Advisory Committee on Rural Health & Human Services, 2012; Rural Poverty Research Institute, 2008). Historically, social services and physical and mental health programs that constitute “wrap-around services” for Head Start families have been less accessible in rural, compared with urban areas (Chertow, 1968). Similarly, the availability of (and/or demand for) alternative, formal child-care options tend to be more limited in rural settings, with research showing that children from rural environments are significantly more likely to receive care from relatives and less likely to be enrolled in center-based care than their urban peers (Miller & Votruba-Drzal, 2013; Swenson, 2008). Indeed, the original Head Start Impact Study report postulates that the “difficulties that children and families in non-urban communities have in getting comprehensive services and in finding quality early care and education for their children” may be responsible for impact variation across these settings (Puma et al., 2010a, p. 9-9).

There is also evidence to suggest that Head Start centers may serve different types of children and families in urban versus rural communities, or, said another way, the children and families in rural and urban Head Starts may differ from one another. Recent research has found, for example, that urban Head Start families show higher levels of educational engagement but lower levels of parent-child attachment than rural Head Start families (Bender, Fedor, & Carlson, 2011; Keys, 2014). Similarly, national data from the Early

Childhood Longitudinal Study Birth Cohort suggest that the lower pre-academic skills of children living in both rural and highly urban settings upon kindergarten entry may be partially explained by higher levels of family socioeconomic adversity in these settings as compared with those in small urban or suburban ones (Grace et al., 2006; Miller & Votruba-Drzal, 2013).

Despite the fact that urban and rural communities, centers, and families face substantially different strengths and challenges, research on the degree to which Head Start and other ECCE program impacts may vary based on urbanicity remains limited and inconclusive. In the 1960's, the Ohio-Westinghouse study found non-experimental impacts of Head Start that were twice as large in urban centers with high concentrations of black children relative to those in the full sample (Smith & Bissel, 1970). Results from the Head Start Impact Study final report, on the other hand, showed stronger and more sustained impacts on language and literacy for three-year-old children from rural communities as compared to three-year olds in urban communities (Puma et al., 2010). Importantly, neither of these studies took into account the ways that additional community, center, family, and individual characteristics may have explained this impact variation. Additional research has found that what appears to be an urban-rural gap in Head Start classroom quality may actually be driven by community socioeconomic disadvantage (Resnick & Zill, 2000), a finding that was supported by recent evidence showing lower levels of material and relational classroom quality in high-poverty neighborhoods (McCoy et al., 2015). Given that higher levels of Head Start program quality have been linked with better academic outcomes for children (Bryant, Burchinal, Lau, & Sparling, 1994), understanding the relationships between urbanicity and the resources, interactions, and instruction that are taking place in classroom settings is a particularly important area of needed research.

Exploring Ecological Sources of Impact Variation

Despite limited understanding of contextual-level predictors of treatment impact variation in Head Start, exploration of moderation in ECCE research is far from novel. A growing body of research has shown consistent evidence for Head Start's differential effectiveness across a number of individual and family characteristics, including stronger impacts for children from families facing high levels of socioeconomic adversity (Cooper & Lanza, 2014; Lee, 2011), dual language learners (Bloom & Weiland, 2015; Puma et al., 2010a), and children with low levels of baseline skills (Lee, 2011; Puma et al., 2010a). For other individual-level characteristics, evidence is more mixed, with studies alternately finding stronger versus weaker impacts of Head Start for children of depressed mothers (Puma et al., 2010a; Robinson & Emde, 2004) and racial/ethnic minority children (Garces, Thomas, & Currie, 2000; Puma et al., 2010a). Although this research has been critical for moving the ECCE field forward, it provides limited information to policy makers aiming to improve program services at scale. Perhaps a key constraint to this area of research is the fact that much research on Head Start and ECCE programs has been conducted in urban settings, limiting researchers' ability to examine cross-site or cross-neighborhood variability in impacts. As a result, almost nothing is known about the degree to which Head Start may be more or less effective in diverse neighborhoods across the country.

In the present paper, we take an ecological approach to understanding urban-rural differences by considering the ways that urbanicity relates to a number of key contextual characteristics that may ultimately drive impact variation across settings. We build on the neighborhood research presented above by considering the ways that urbanicity is inter-related with a number of key *community characteristics* for low-income families, including neighborhood poverty, rates of crime, racial/ethnic distribution, availability of alternative public preschool options, and concentration of neighborhood resources. We also draw from the descriptive research on urban and rural ECCE programs by considering the ways that urban-rural differences in program impacts may be driven by *center characteristics*, including centers' average teacher education levels, provision of transportation services, difficulty in recruiting teachers, full-versus part-time care options, provision of family and child services, and average levels of classroom resources, relational quality, and instructional activities. Finally, we consider the role of *compositional effects*, or the ways that certain groups of children and families concentrate differently across urban and rural environments. In particular, we focus on characteristics found in previous research to predict differential effectiveness of programs, including children's baseline skill levels, demographic characteristics (e.g., age, gender, race/ethnicity), and home language, as well as mothers' marital status, education, immigration status, and depressive symptoms.

The Present Study

Given the lack of data on Head Start neighborhoods nationally, the goals of this research are descriptive and hypothesis-generating. The first aim of the present study is to provide a descriptive "snapshot" of the community environments in which Head Start programs are based using data from the national Head Start Impact Study (HSIS). In particular, we focus on whether and how Head Start families and centers are non-randomly distributed across urban and rural settings. Based on previous research, we hypothesize that urban Head Start centers will generally show higher levels of socioeconomic adversity, higher availability of resources and services, and more racial/ethnic and linguistic diversity than their rural counterparts.

The second aim of this study is to test the ways that centers' location in urban versus rural environments may predict differential impacts of Head Start on children's early language and literacy outcomes. Building on this, our third aim is to understand the degree to which additional contextual characteristics – including neighborhood, center, and individual compositional characteristics – may account for any observed impact variation attributed to urbanicity in aim two. To achieve these aims, we apply new multi-level modeling approaches (Bloom & Weiland, 2015) to rich individual, family, and contextual data from the HSIS and publicly available census sources to provide a new, ecological approach to program impact evaluation.

In this study, we focus specifically on children's early language processing (i.e., receptive vocabulary and oral comprehension) and pre-literacy outcomes (i.e., early reading and early writing) due to their importance within the Head Start Outcomes Framework (Office of Head Start, 2010), as well as their relative underrepresentation in neighborhood research. We consider these outcomes separately given previous research showing unique main and

moderating effects of urbanicity on each, including evidence from the original Head Start Impact Study (e.g., Grace et al., 2006; Puma et al., 2010a). In particular, previous research in the K-3 educational system has shown urban-rural “gaps” in the emphasis that is placed on various aspects of language and literacy, with urban districts focusing more heavily on core instruction related to vocabulary, phonics, and reading, and less heavily on comprehension and writing than their rural counterparts (Teale, Paciga, & Hoffman, 2007). Based on this work and the research highlighted above, we hypothesize that the impacts of Head Start on children’s early vocabulary and reading outcomes will be largest for urban centers, and on comprehension and writing will be largest for rural centers. We also hypothesize that these differences will attenuate when accounting for important contextual characteristics that covary with urbanicity, including neighborhood resources, centers’ levels of quality and use of literacy-related instructional activities, and centers’ inclusion of high percentages of sociodemographically diverse children.

Method

Sample

The original HSIS was designed to be nationally representative of 3- and 4-year-olds attending Head Start programs in the United States and included 4,440 children across 22 states. Children who applied to one of 351 traditionally oversubscribed Head Start programs across 81 Head Start grantees were randomly assigned either to receive an invitation to participate in Head Start services ($n = 2,644$) or to the control group ($n = 1,796$). Following randomization, these children enrolled in a total of 1,632 classrooms across 930 Head Start and non-Head Start programs, including 42 family child care homes; the remaining children were cared for exclusively in their own homes by a parent or other adult, or were missing information about their ECCE setting. In total, 2,124 children enrolled in Head Start, 638 children enrolled in a non-Head Start center-based program, and 204 children attended family child care.

The present study’s analytic sample differs from the original, nationally representative HSIS sample in several important ways. First, the current study excludes 845 children with missing Spring 2003 outcome data. Second, in order to preserve the experimental contrast in each random assignment site, we also excluded 92 children from 34 centers that do not have at least one child assigned to the treatment condition and one to control. Third, we did not use the original study’s sampling weights due to the weights’ non-representativeness within our more limited sample, as well as their computational incompatibility with our analytic approach (see below). Thus, the final, unweighted analytic sample for this study includes 3,503 children from 317 random-assignment sites. The average age of this sample was 4.04 years old. Children (51% female) were racially diverse: 31% were Black and 35% were Hispanic. Approximately 37% of children in the sample had a mother who had not earned a high school diploma and 55% of children had an unmarried, single mother. Parents reported that English was spoken in the home for 72% of children, and 17% of children had parents who were recent immigrants. (See Table 1 for additional descriptive data.)

Procedure

Recruitment and random assignment—During the summer of 2002, first-time 3- and 4-year old applicants to one of the participating Head Start centers were randomly assigned to either receive the offer to participate in Head Start or to the control group. Although members of the control group were not offered a slot in Head Start, they could enroll in other types of center-based care, family child care, or in-home care. Prior to random assignment, parents received information about the HSIS, including procedures, potential benefits, and study incentives. Following random assignment, study personnel met with parents in groups and individually to further explain the study. Parents were asked to provide informed consent for the duration of the study, which included allowing their child to participate and be assessed, and permitting the researchers to contact the child's teacher (Puma et al., 2010b). Parents and teachers received small cash incentives for their participation in the study (i.e., completion of interviews and questionnaires). Center directors received a cash incentive for permitting classroom observations (Puma et al., 2010b).

Data collection—Data collection began in the fall of 2002 when children entered Head Start for the first time and continued throughout the school year. During the fall of 2002 and spring of 2003, trained data collectors visited children at their ECCE program or at home (for children not enrolled in an ECCE program) to administer language and literacy assessments. Children's primary caregivers also completed a questionnaire about their child and family during the fall of 2002 and the spring of 2003 (Puma et al., 2010a). During the winter and spring of 2003, trained data collectors visited treatment and control children's primary ECCE program to conduct director interviews. Interviews were conducted in both Head Start and non-Head Start center-based programs, and include questions regarding staff qualifications, services provided, and program structure.

Geocoding—In addition to the use of publicly available HSIS data, the present study utilized an additional set of restricted information that was provided to the [center name redacted] as part of a broader agreement with the Administration for Children and Families. In particular, geocodes (i.e., latitude and longitude) of each random-assignment center were provided for use by the [name redacted]. Geocodes of children's homes and post-random assignment ECCE center of attendance were not provided. Geocodes for each random assignment center were coded and linked to three neighborhood boundaries – surrounding census tracts, zip codes, and counties – using ArcGIS software (Version 10.1; ESRI, 2011). These boundaries were chosen to match with the level of measurement of the various data sources used for capturing neighborhood characteristics.

Measures

Child outcomes—Children's early English language and literacy outcomes included measures of children's early language processing (i.e., receptive vocabulary and oral comprehension), as well as two dimensions of pre-literacy skills (i.e., early reading and early writing). Children's *receptive vocabulary* was captured using the Peabody Picture Vocabulary Test (PPVT; Dunn, Dunn, & Dunn, 1997). In the PPVT, the child looks at four pictures and is instructed to point to the picture that best matched the word spoken by the

assessor. Previous work using the PPVT has shown it to have adequate internal consistency, split-half reliability, test-retest reliability, and validity (Dunn et al., 1997; Puma et al., 2010b). Item response theory was used to develop a shorter version of the PPVT that was used in the HSIS, with scoring representing children's placement on a continuous scale defined by the IRT item difficulty, discrimination, and guessing behavioral parameters (see Puma et al., 2010b). *Oral comprehension* was measured using the Woodcock-Johnson Oral Comprehension test (WJOC; Woodcock, McGrew, & Mather, 2001), which measures the child's ability to comprehend a short, spoken passage and to provide a missing word based on syntactic and semantic clues. Previous work has shown its test-retest reliability and internal consistency to be strong (Woodcock et al., 2001).

Pre-literacy skills were measured using the Woodcock-Johnson tests of Letter-Word Identification (WJLW) and Spelling (WJS; Woodcock et al., 2001). *Early reading* was captured with the WJLW, which measures a child's skills in identifying letters and words as they appear in the test easel. The test includes both receptive ("point to the [x]") and expressive ("name this letter/word") components. In addition, children's *early writing* was captured by the WJS, in which pre-writing skills are measured through tasks such as drawing lines and copying letters. As the items progress in difficulty, the child is asked to write specific upper and lower cases of the alphabet and specific words. Both the WJLW and WJS have been used extensively in previous work assessing children's pre-literacy, with ample evidence for their reliability and validity (Puma et al., 2010b; Woodcock et al., 2001).

All Woodcock-Johnson scores were standardized to represent a W-ability score, which represents a linear item response theory score obtained from the publisher based on a Rausch model. Means, standard deviations, and ranges for the study sample can be found in Table 1. For more details on all outcome measures and approaches to standardization, see Puma et al., 2010b.

Urbanicity—Data on Head Start centers' urbanicity were taken from the United States Decennial Census for the year 2000. Urbanicity was represented by the proportion of households within the random assignment center's census tract that lived in an urbanized area (UA) or an urban cluster (UC) in the year 2000. UAs were defined as areas with population densities of more than 1000 people per square mile, whereas UCs were defined as areas surrounding UAs with population densities of at least 400 people per square mile. Although most census tracts included either 100% of households falling in a UC or UA (i.e., fully urban tracts) or 0% of households falling in a UC or UA (i.e., fully rural tracts), a small but substantial number of census tracts included a mixture of UC/UA and non-UC/non-UA households (i.e., mixed urbanicity or suburban tracts). To preserve this heterogeneity, the urbanicity variable used in the present study is a continuous variable ranging from 0 to 1 rather than dichotomous variable representing rural versus urban.

Neighborhood covariates—Covariates representing the neighborhoods surrounding Head Start random assignment centers were obtained from several publicly available data sources, including the U.S. Census, Federal Bureau of Investigation (FBI), and the U.S. Department of Education (DOE). Characteristics potentially associated with urbanicity were our focus; hence, we sought measures of 1) poverty; 2) racial/ethnic composition; 3) crime

rates; 4) the number of alternative early education and care arrangements; and 5) the availability of social and commercial neighborhood resources.

Head Start neighborhood poverty was represented by the proportion of households within the random assignment center's census tract that fell below the federal poverty line in the year 2000 (two years prior to data collection). In addition, neighborhood racial/ethnic minority concentration was represented by the proportion of individuals within the Census tract that did not identify as non-Hispanic white. These data were taken directly from the U.S. Census Bureau's decennial census (2000).

The total number of violent and non-violent crimes per 1000 people was obtained at the county level by request from the Criminal Justice Information Services division of the FBI for the year 2002.

The number of DOE-funded schools that include a preschool program within the random assignment zip code was taken from the National Center for Education Statistics (NCES) Common Core of Data for the 2002–2003 school year (U.S. Department of Education, 2004). As this variable does not include privately funded center- or home-based preschool programs, public programs that are funded through non-DOE federal, state, or local sources (including other Head Start centers), or DOE-funded preschool programs that are not housed in public school buildings (a group estimated to make up about one half of all DOE-funded preschool programs in 2001; Clifford et al., 2005), we rely on this information as a proxy for alternative care availability rather than a comprehensive measure. For analysis, the number of DOE preschool programs per 100 children under the age of 5 was calculated using 2000 U.S. Census sample data.

Finally, the availability of social and commercial resources in Head Start center zip codes was captured using data from the Zip Code Business Pattern section of the U.S. Census from 2002. The Zip Code Business Pattern data categorize all businesses and organizations with a payroll into over 1,000 North American Industry Classification System (NAICS) codes. Using the NAICS codes, we identified 176 types of establishments likely present in Head Start random assignment center zip codes, including grocery stores, schools, churches, museums, clothing stores, legal offices, and theatres. (A full list of all establishments available upon request from the authors.) The total resources variable reflects the total number of these resources present in Head Start random assignment center zip codes.

Center covariates—Head Start random assignment center characteristics were reported by program directors and teachers, and directly observed by trained data collectors in the spring of 2003¹. Again, covariates were considered that were theorized to be associated with urbanicity, the central focus of our research questions. Specifically, directors' reports of 1) the proportion of teachers who have obtained a bachelor's degree; 2) the number of vacant teaching positions; 3) the provision of full-day care; 4) availability of local transportation;

¹All program director surveys were conducted in the settings in which children attended ECCE, which could not be directly matched to centers of random assignment using publicly available data. A multi-stage process that examined the concentration of treatment and control children in each center of attendance was used to create a match between IDs for attendance and random assignment. For details of this match process, please contact the first author.

and 5) a set of other child and family services (e.g., counseling, parenting education; a total of 20 services) are included in the current analyses. In addition, teachers reported on the frequency with which they used literacy-related instructional techniques (e.g., book reading, practicing letter sounds; a total of 12 techniques). Finally, classroom quality was captured using the Early Childhood Environment Rating Scale–Revised (ECERS-R) and the Arnett Scale of Lead Teacher Behavior (Arnett, 1989; Harms, Clifford, & Cryer, 1998). Items from the ECERS-R and Arnett were grouped into three domains: material and spatial quality, positive teacher-student interactions, and negative teacher-student interactions—based on factor analysis (see Connors, Friedman-Krauss, Morris, & Jones, 2015, for details). The number of child and family services provided by the program, the frequency of literacy-related activities, and the items comprising the three classroom quality domains were each consolidated into an index score with a possible range from 0 to 1. Literacy activities and classroom quality variables were averaged across classrooms within random assignment centers to represent center-level characteristics.

Child and family composition covariates—Given the different populations of families attending rural and urban centers, a set of child- and family-level covariates was also used for analyses, including child gender, child age, child cohort (an indicator for whether the child was two years vs. one year away from Kindergarten entry), child race (an indicator variable for children who were Hispanic and an indicator for children who were black), maternal education (an indicator for less than high school education), home language (an indicator for English), recent immigrant status (an indicator for living in the United States for less than 10 years), and mother’s marital status (an indicator for single mother). These characteristics were reported by the primary caregiver in the fall of 2002. Mothers’ depressive symptoms were also measured at this time using the Center for Epidemiologic Studies–Depression scale (CES-D; Seligman, 1993).

Analytic Plan

—To address our first study aim, we use descriptive analysis and bivariate correlations to a) characterize the neighborhoods of Head Start centers and b) examine the way in which urbanicity of the Head Start neighborhood is related to other characteristics of individuals, families, centers, and neighborhoods.

To examine the degree to which Head Start is more or less effective in urban versus rural environments (our second aim), we build upon new approaches to quantifying impact variation in randomized controlled trials (Bloom, Porter, & Weiss, 2013; Raudenbush & Bloom, 2015; Raudenbush, Reardon, & Nomi, 2012). In particular, we build from a model developed by Bloom and Weiland (2015) using the HSIS data that includes individual children at level 1 and the Head Start centers at which they initially sought out care (i.e., their random assignment center) at level 2. By nesting children in their random assignment center rather than their center of attendance, we were able to preserve the experimental integrity of the study design and estimate the causal impacts of random assignment to Head Start on children’s outcomes using an intent to treat approach. In our sample, 18 percent of those assigned to treatment did not end up in a Head Start center (i.e., were “no shows”) and 14 percent of those assigned to the control group did end up in a Head Start center (i.e., were

“crossovers”). Of those who were assigned to Head Start and took up the offer (i.e., “treatment compliers”), we estimate that 92 percent attended Head Start in their center of random assignment, whereas 8 percent attended somewhere else.

Our model for aim two differs from traditional methodological approaches in several ways. First, unlike traditional multi-level model with random intercepts, our model includes fixed effects for Head Start random assignment center and random slopes for the effects of treatment at level 1. The use of fixed effects for centers eliminates biases resulting from systematic and non-systematic differences across Head Start centers, whereas the modeling of separate residual variances for treatment and control conditions accounts for any potential effects of random assignment on the distribution of outcome variances (Bloom & Weiland, 2015).² Second, unlike traditional approaches to moderation, our approach conceptualizes urbanicity as a predictor of the Head Start treatment effect. Mechanically, this means that we include urbanicity at level 2 as a predictor of the level 1 slope for treatment, thereby creating a cross-level interaction term. Specifically, we estimate the following model:

Level 1 (individual)

$$Y_{ij} = \alpha_j + \beta_{1j} * Treat_{ij} + \sum_{k=2}^K \beta_{kj} * X_{kij} + e_{ij}$$

Level 2 (center/neighborhood):

$$\begin{aligned} \alpha_j &= \alpha_j \\ \beta_{1j} &= \gamma_{10} + \gamma_{11} Urban_j + r_j \end{aligned}$$

where:

Y_{ij} = the value of the outcome for individual i from Head Start random assignment center j ,

$Treat_{ij}$ = 1 if individual i from center j was randomized to Head Start and 0 otherwise,

X_{kij} = the value of baseline characteristic k for individual i from center j , including pre-test scores for the relevant outcome of interest

α_j = the mean control group outcome for center j (a parameter that is fixed for each center),

β_{1j} = the mean Head Start effect for center j ,

β_{kj} = the mean non-causal “effect” of covariate k for center j ,

γ_{10} = the cross-site grand mean effect of random assignment to Head Start,

²A set of models with random intercepts instead of fixed intercepts was run as a sensitivity analysis. These results were highly similar to the results from the fixed intercept models and are therefore not included in this paper.

γ_{11} = the coefficient for the non-causal “effect” of center neighborhood urbanicity on the mean Head Start effect,

e_{ij} = a random error that varies across individuals with a zero mean and variances σ_T^2 and σ_C^2 for treatment and control group members, respectively,

r_j = a random error that varies across centers with mean 0 and variance τ^2

The third aim of this study is to determine whether observed differences in Head Start impacts across urban versus rural environments may be attributable to covarying neighborhood, center, and compositional characteristics. To test this question, we conduct an additional set of analyses that control for these characteristics in several ways. First, we estimate a model for each outcome that, in addition to the parameters noted above for estimates of the effect of urbanicity on the treatment impact, also includes an additional set of potentially confounding neighborhood covariates at level 2 as predictors of the level 1 coefficient for treatment. Second, we estimate a model for each outcome that also includes a set of potentially confounding center covariates at level 2 as predictors of the level 1 coefficient for treatment. Third, we estimate a model for each outcome that also includes a set of potentially confounding child and family covariates at level 1 as both main effects and as interactions with treatment. Due to the presence of fixed intercepts in our equation, this particular model allows us to adjust for the effect of these potential confounds on the treatment impact within each center, thereby controlling for differential impacts based on the sociodemographic composition of individuals and families within these centers (Bloom, personal communication). Finally, we examine all of the above potentially confounding characteristics in a single model for each outcome.

Across all of these additional sets of analyses, our goal is to determine the robustness of urbanicity as a predictor of Head Start’s impact to the inclusion of several sets of potentially confounding characteristics. Similar to previous work (e.g., Bumgarner & Brooks-Gunn, 2015; Zhang, 2006), we evaluate the degree to which the γ_{11} coefficient (the effect of urbanicity on the level 1 slope for treatment) was *attenuated* in terms of its significance and magnitude when accounting for these additional characteristics as predictors of the treatment effect. Because we cannot test whether the γ_{11} coefficient is significantly different across non-nested models, we instead examine whether there is a qualitatively meaningful attenuation of the γ_{11} coefficient that would suggest that the role of urbanicity in predicting differential impacts of Head Start may be confounded. For example, if a previously significant γ_{11} coefficient is attenuated to zero or close to zero when accounting for other neighborhood characteristics, we would conclude that any effect that urbanicity had on Head Start’s effectiveness to impact that particular outcome was likely attributable to some other aspect of the neighborhood (e.g., crime rates, poverty, resource availability) that was correlated with urbanicity. Although such an approach does not provide a definitive, causal answer to our question of what is driving impact variation, it provides initial evidence that can be explored using more precise methods in future work.

Across all analyses, we use a cutoff of $\alpha < .10$ to determine statistical significance. This choice reflects the exploratory nature of our study and is in line with previous work using the HSIS dataset (e.g., Bloom & Weiland, 2015; Puma et al., 2010a; 2010b). In addition, the

original HSIS was designed to detect the statistical significance of the main effect of Head Start on children's outcomes, but was not explicitly powered to detect impact variation. Indeed, power analyses suggest that, holding sample sizes constant, the minimum detectable effect size of a main effect is approximately half that of a center-level predictor of impact variation (Bloom, Raudenbush, & Reardon, 2014; Spybrook, 2014). As we cannot control the sample size of the HSIS, we instead use a more liberal cutoff for statistical significance that will allow us to detect subgroup impacts that are of the same magnitude of those observed in the full sample. For primary analyses, all variables (with the exception of treatment) were grand-mean centered for analysis, allowing impact estimates to represent effects for the average child. In the case that significant differences by urbanicity were observed, we also generated *post hoc* estimates of the effects of random assignment to Head Start for "typical" children within urban, rural, and mixed (i.e., half urban, half rural) communities by re-centering urbanicity to reflect the particular community of interest.

Missing Data

As is noted above, our analytic sample excludes a total of 845 children with missing data on their spring outcome scores (von Hippel, 2007). Rates of missing predictor data in our final analytic sample are shown in Table 1. All missing predictor variable data were addressed through single replication of a multiple imputation model conducted separately for level 1 and level 2 variables. Specifically, we use all valid baseline and follow-up information to impute missing predictor variable values (Little & Rubin, 2014). This approach was used to minimize complexity yet maintain use of the larger sample. Previous analyses in this sample using the same approach have shown this strategy to be justified based on a lack of evidence for: 1) bias due to sample attrition, 2) variation in results based on the presence or absence of covariates, and 3) differences in results across various imputation techniques (Bloom & Weiland, 2015).

Results

Descriptive Findings

Table 1 highlights the means, standard deviations, and ranges of the neighborhood, center, family, and individual characteristics used as primary study variables in the present analyses. These results show high levels of variability in the characteristics of neighborhoods surrounding Head Start centers, the resources of the centers themselves, and the individuals and families they serve. In particular, we find that approximately 59% of centers in the HSIS were located in fully urban neighborhoods, 31% were located in fully rural neighborhoods, and 11% were located in neighborhoods containing both urban and rural households (i.e., suburban or mixed urbanicity environments). Table 1 highlights the bivariate correlations between urbanicity and all study variables. These results show that Head Start neighborhood urbanicity is significantly and positively correlated with neighborhood poverty, crime, ethnic minority concentration, and resource availability, as well as teacher education within Head Start centers. Urbanicity is also negatively correlated with centers' provision of transportation and child/family services. At the individual level, the characteristics that were most associated with being in an urban Head Start center were

children's Hispanic ethnicity, speaking a language other than English in the home, and maternal depressive symptoms.

Results of Multi-level Regression Analyses

Results of our basic model examining the role of neighborhood urbanicity in predicting differential impacts of Head Start on child outcomes are shown in the first set of columns of Tables 2–5. These results suggest that urbanicity is a marginally significant positive predictor of the *impact* of Head Start assignment on children's receptive vocabulary (PPVT), $b = 3.64$ ($SE = 2.15$), $p < .10$, and a significant negative predictor of the impact of Head Start assignment on children's oral comprehension skills, $b = -2.03$ ($SE = 0.94$), $p < .05$.

Urbanicity did not predict the impact of Head Start assignment on either letter word identification or spelling. As is shown in the first panel of Figure 1, this moderation effect can be interpreted as follows: the positive impact of random assignment to Head Start on children's receptive vocabulary skills was larger for centers located in urban neighborhoods (defined by an urbanicity score of 1) relative to those in rural communities (defined by an urbanicity score of 0). Specifically, *post hoc* analyses suggest assignment to Head Start produced small, positive, and statistically significant gains in children's receptive vocabulary skills in urban, $b = 6.44$ ($SE = 1.18$), $p < .01$, $d = 0.16$, and mixed (50% urban, 50% rural) neighborhoods, $b = 4.62$ ($SE = 1.06$), $p < .01$, $d = 0.12$, but did not have a statistically detectable impact in rural communities, $b = 2.80$ ($SE = 1.78$), *ns*, $d = 0.07$. For oral comprehension skills, on the other hand, the effect of random assignment to Head Start was larger in rural relative to urban communities (see Figure 2). In particular, Head Start had no statistically detectable impact in urban, $b = -0.33$ ($SE = 0.51$), *ns*, $d = -0.02$, or mixed neighborhoods, $b = 0.69$ ($SE = 0.46$), *ns*, $d = 0.04$, but did have a small, positive, and statistically significant effect in rural communities, $b = 1.71$ ($SE = 0.78$), $p < .05$, $d = 0.10$.

To explore the degree to which these findings were robust to the inclusion of other potential contextual predictors of differential impact related to urbanicity (aim 3), we next tested a series of models that accounted for additional, potentially confounding characteristics. Below and in the latter columns of Tables 2–5 we report the results of the models that separately assess the role of neighborhood characteristics, center characteristics, and compositional characteristics. Results of the models including all of these characteristics simultaneously were largely consistent with, though slightly attenuated relative to these more targeted models and are therefore not shown in detail. (For results of these full models, please contact the first author.) For receptive vocabulary, urbanicity was not a statistically significant predictor of random assignment impacts in any of these additional models (see Figure 1). The magnitude of the coefficient for urbanicity, however, remained relatively consistent across models that accounted for neighborhood and center characteristics, indicating that the lack of significance in these models may have been attributable to reduced statistical power rather than any meaningful evidence of confounding effects. The magnitude of the coefficient for urbanicity did, however, attenuate by 35% in the model accounting for the characteristics of children and families, $b = 2.38$ ($SE = 2.24$), *ns*. Additional analyses including each of the individual and family characteristics as predictors of the treatment coefficient separately in the basic model suggested that the coefficient for urbanicity was attenuated to $b < 2.50$ with the inclusion of any of the following: child

Hispanic ethnicity, home language, or maternal immigrant status. The coefficient for urbanicity remained statistically significant (at $p < .10$) and above $b = 3.50$ when accounting for any of the other child/family characteristics on their own as predictors of the treatment coefficient.

Additional analyses exploring the robustness of the variability in impacts for oral comprehension revealed that the magnitude of the coefficient for urbanicity remained relatively unchanged across models accounting for neighborhood and individual/family composition characteristics (see Figure 2). The urbanicity coefficient in the model accounting for center characteristics, however, was attenuated by about 33% to $b = -1.37$ ($SE = 1.02$), *ns*. Additional analyses revealed that this attenuation was primarily driven by the inclusion of centers' provision of transportation. The inclusion of transportation alone led to a coefficient of $b = -1.50$ ($SE = 0.97$), *ns*, whereas the models including the other center characteristics showed urbanicity coefficients of $b < -1.90$ that were also statistically significant at $p < .05$.

Discussion

The primary aim of the present study was to understand whether Head Start's impacts on children's early gains in language and literacy skills across the preschool year were similar or different for centers characterized by varying levels of urbanicity. The results of our analyses suggest that random assignment to Head Start led to small but significant gains in receptive vocabulary skills relative to the control group in urban and mixed communities, but that it was not effective in rural areas. For oral comprehension skills, however, the impacts of Head Start were small and positive in fully rural environments, whereas they were non-significant in both urban and mixed environments. Importantly, these results were not robust to the inclusion of features of centers and the children they served, suggesting that other characteristics of children's environments – including the composition of non-English-speaking children and families within the center and centers' provision of transportation – likely underlie these differential impacts. These results complement and extend previous work showing stronger impacts of Head Start on language and literacy outcomes for children who come from socioeconomically at-risk families, Spanish-speaking households, and environments with moderate levels of parental pre-academic stimulation (Bloom & Weiland, 2015; Cooper & Lanza, 2014; Miller, Farkas, Vandell, & Duncan, 2014).

To contextualize these distinct patterns, it is first important to understand the ways in which receptive vocabulary and oral comprehension are different from one another conceptually, as well as how they are emphasized (or not) across diverse Head Start environments. Although both represent children's listening skills and vocabulary knowledge, receptive vocabulary and oral comprehension skills may emerge as the result of different types of environmental inputs and instructional techniques. As a basic measure of receptive vocabulary, the PPVT is thought to be sensitive to explicit forms of English vocabulary instruction that emphasize declarative knowledge about words, their definitions, and their usage (Nagy & Scott, 2000). The development of broader oral comprehension skills, on the other hand, requires a more complex set of skills that combine declarative knowledge with the ability to contextualize words and understand their underlying meaning (Beck &

McKeown, 2007; Hart & Risley, 1995; Lonigan, Anthony, Bloomfield, Dyer, & Samwel, 1999). As a result, research has shown that broader exposure to linguistically diverse and complex classroom environments (e.g., those that include child-directed conversations during free play, those that include “decontextualized” discussions of objects and events that are not physically present, those that explicitly relate words to real-life situations) is particularly important for children’s development of the decoding, analytic, and reasoning skills that are measured in the WJ Oral Comprehension scale (Dickinson, 2001).

Given the etiology of these skills and Head Start’s focus on explicit vocabulary instruction (Hindman & Wasik, 2013), it is not surprising that Head Start demonstrated positive effects on children’s receptive vocabulary scores across both urban and rural environments. In particular, our results suggest that the larger and significant effects of Head Start on children’s PPVT scores in urban environments may have been driven by particularly high concentrations of Spanish-speaking dual language learners (DLLs) in these urban contexts (Gibson & Jung, 2006). The provision of culturally and linguistically responsive, bilingual (and likely vocabulary-focused) instruction and services to DLLs – which compose approximately one quarter of Head Start participants nationally (Administration for Children and Families, 2013) – is central to Head Start’s mission, and may account for current research showing significantly greater improvements for DLLs’ English vocabulary in Head Start compared with non-DLL’s (Bitler, Hoynes, & Domina, 2014; Feller, Grindal, Miratrix, & Page, under review; Oh, Yoshikawa, & Willet, 2015). In particular, our findings align with recent work suggesting that Head Start is more effective than alternative options in compensating for DLL’s low English vocabulary skills upon preschool entry (Bloom & Weiland, 2015; Cooper & Lanza, 2015). These findings also support a growing body of research highlighting the utility of linguistically sensitive ECCE programming for promoting DLL and immigrant children’s early language and literacy skills (Farver, Lonigan, & Eppe, 2009; Gormley, 2008; Magnuson, Lahaie, & Waldfogel, 2006).

By contrast to findings on receptive vocabulary, impacts of Head Start on the WJ Oral Comprehension scale were positive in rural environments and null in urban environments, a difference that was attenuated when accounting for centers’ provision of transportation services. In general, the smaller impact of Head Start on oral comprehension in urban versus rural environments is likely not simply a function of urban centers’ lower provision of transportation (which may not be necessary where children live within walking distance of their schools or public transportation), but rather unmeasured center characteristics that may covary with transportation. Although we attempted to account for a number of theoretically relevant center, classroom, and teacher characteristics in the present study, the measures used to represent these processes were relatively coarse and likely did not capture the full breadth of possible causal mechanisms. It is possible, for example, that relative to their local competition, low-resourced urban Head Start centers are unable to provide high doses of the diverse, sophisticated, and decontextualized linguistic environments that are necessary to boost children’s oral comprehension skills. In rural areas, on the other hand, Head Start centers may be better equipped to support these skills, either due to higher attendance rates, better curricular approaches, more funding for language-related professional development, differences in parental expectations and involvement in literacy, and/or lower-quality language environments in counterfactual group contexts. Future research capturing a greater

diversity of these ecological characteristics using more nuanced approaches to measurement is needed to disentangle these urban-rural processes.

In general, these findings highlight several important considerations for understanding ECCE program effectiveness across settings and child outcomes. First, although we observed differences in Head Start impacts for urban and rural environments in the domains of receptive vocabulary and oral comprehension, we also found evidence that this variation was partially – though not fully – driven by observed differences in characteristics of the centers and individuals who “selected” into these contexts. These findings support previous empirical theoretical work on neighborhoods and children’s development, and highlight the importance of taking an ecological approach to understanding the unique roles that neighborhoods, educational settings, families, and peers play in shaping children’s growth in the context of overlapping risk/advantage (Aikens & Barbarin, 2008). Second, the original HSIS report identified positive and significant overall benefits of Head Start on children’s receptive vocabulary, yet null average effects on oral comprehension (Puma et al., 2010a). When decomposing the full sample into urban and rural environments, the present study contrasts these original results by showing null effects of Head Start on vocabulary and small but significant positive impacts for oral comprehension in rural communities. These results reinforce the importance of considering particular subpopulations of children, families, centers, and communities for identifying key questions of “what works” and “for whom,” regardless of the statistical significance of overall program impacts. Finally, the lack of significant urban-rural impact variation for early reading and early writing suggests that unlike for early language skills, Head Start may be equally well equipped to promote children’s literacy-related pre-academic knowledge across urban versus rural community settings. Future research is needed to explore the degree to which contextual characteristics other than urbanicity may serve as more salient predictors of impact variation for these and other developmental domains.

Limitations and Future Directions

Although this study has numerous strengths – including its novel exploration of urban-rural differences, its focus on a multi-state sample, and its use of contemporary analytic approaches with pre-test controls – it is not without limitation. First, although we capitalize on the experimental nature of the Head Start Impact Study, our primary predictor of impact variation – urbanicity – was not randomly assigned. For this reason, we cannot attribute causality to the difference in impacts across urban and rural environments, nor can we fully disentangle the distinct effects of specific environmental features. Second, there are a number of social, structural, and relational characteristics of urban environments that are not captured in the present study, making it difficult to fully explain all processes that may be responsible for observed differences. In particular, data on attendance rates and the quality of the instructional and linguistic environments available to children at both Head Start and counterfactual centers would be useful for contextualizing our results further. Third, although our use of U.S. Census data allows us to evaluate urbanicity continuously, it does not capture potentially important but more qualitative gradations in urban structures (e.g., distinguishing between large metropolitan cities, smaller cities or towns, suburban environments, and more rural communities). Fourth, although the Head Start Impact Study

was designed as a nationally representative sample of Head Start programs, the complexity of our analytic approach precluded our use of sampling weights and necessitated the exclusion of specific portions of the sample. This, combined with the fact that these data were collected more than a decade ago, limits the generalizability of these findings to the current national population of Head Start programs.

In the future, research using more current samples with rich contextual data will be useful for further exploring the multiple environmental sources of ECCE program impact for low-income children nationally. In particular, large-scale randomized controlled trials of ECCE programming are needed that simultaneously capture detailed and actionable data at multiple ecological levels, including information on children's previous learning experiences in and outside of the home, how caregivers make decisions about childcare, the types of instructional practices used by teachers in the classroom, and state and local childcare policy. To ensure the field's ability to answer questions of "why" and "for whom," such studies must be sufficiently powered not only to detect main effects, but also to further explore questions of mediation and moderation at multiple levels (Spybrook, 2014). Finally, future research is also needed to explore the durability of these findings beyond the preschool year, including the possibility of sustained effects through elementary school, adolescence, and adulthood within particular communities that may be masked by fade-out in others.

Conclusions

Despite these limitations, the results of this study contribute important knowledge to policy-makers and practitioners interested in improving low-income children's language and literacy skills. Results of this study highlight the ways that resources, services, and risk factors related to neighborhoods, centers, and families are non-randomly distributed across urban and rural environments in ways that have implications for Head Start's ability to provide equitably effective programming. As a result, a "one size fits all" approach to Head Start and other ECCE programs is inappropriate for optimizing outcomes for diverse populations children. Whenever possible, Head Start policies, funding, and curricula should reflect this important diversity, allowing flexibility in service provision to meet local needs and emphasize community strengths. In particular, results of this research suggest that curricula and instruction targeting children's receptive vocabulary skills are particularly necessary in rural communities and those serving primarily native English speakers, whereas approaches oriented toward children's oral comprehension may be particularly beneficial in more urban settings and those without transportation services. Future research is needed to identify additional contextual characteristics that support and undermine the efficacy of ECCE programming as a means for improving children's outcomes across program, setting, and time.

Acknowledgments

The research reported here was funded under cooperative agreement #90YR0049/02 with the Administration for Children and Families (ACF) of the U.S. Department of Health and Human Services and was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R305B080019 to New York University as well as by the Eunice Kennedy Shriver National Institute Of Child Health & Human Development of the National Institutes of Health under Award Number F32HD078034. The opinions expressed are those of the

authors and do not represent views of ACF, the U.S. Department of Health and Human Services, the Institute of Education Sciences, the U.S. Department of Education, or the National Institutes of Health.

References

- Administration for Children and Families. Report to congress on dual language learners in Head Start and Early Head Start programs. U.S. Government; 2013.
- Aikens NL, Barbarin O. Socioeconomic differences in reading trajectories: The contribution of family, neighborhood, and school contexts. *Journal of Educational Psychology*. 2008; 100(2):235.
- Amato PR, Zuo J. Rural poverty, urban poverty, and psychological well-being. *The Sociological Quarterly*. 1992; 33(2):229–240.
- Arnett, J. Caregiver interaction scale. Princeton, NJ: Educational Testing Service; 1989.
- Beck IL, McKeown MG. Increasing young low-income children's oral vocabulary repertoires through rich and focused instruction. *The Elementary School Journal*. 2007; 107(3):251–271.
- Bender SL, Fedor MC, Carlson JS. Examining protective factors and risk factors in urban and rural head start Preschoolers. *Journal of Community Psychology*. 2011; 39(8):908–921.
- Bernardy, PM. Doctoral dissertation. 2012. Head Start: Assessing common explanations for the apparent disappearance of initial positive effects. Retrieved from Mason Archival Repository Service
- Bitler, MP.; Hoynes, HW.; Domina, T. Experimental evidence on distributional effects of Head Start (No. w20434). National Bureau of Economic Research; 2014.
- Bloom, HS.; Raudenbush, SW.; Reardon, SF. Using cross-site variation in program effects to study what works for whom, under what conditions and why?. Paper presented at the annual conference of the Society for Research on Educational Effectiveness; Washington, DC. 2014 Mar.
- Bloom, H.; Weiland, C. 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). 2015. Retrieved from MDRC website: http://www.mdrc.org/sites/default/files/quantifying_variation_in_head_start.pdf
- Bloom, HS.; Porter, KE.; Weiss, MJ. Estimating cross-site impact variation in the presence of heteroscedasticity. Symposium conducted at the Society for Research in Educational Effectiveness; Washington, DC. 2013.
- Boardman JD, Finch BK, Ellison CG, Williams DR, Jackson JS. Neighborhood disadvantage, stress, and drug use among adults. *Journal of Health and Social Behavior*. 2001; 42:151–165. [PubMed: 11467250]
- Brooks-Gunn, J.; Duncan, G.; Aber, JL., editors. Neighborhood poverty, Volume 1: Context and consequences for children. New York: Russell Sage Foundation; 1997.
- Bryant DM, Burchinal M, Lau LB, Sparling JJ. Family and classroom correlates of Head Start children's developmental outcomes. *Early Childhood Research Quarterly*. 1994; 9(3):289–309.
- Bumgarner E, Brooks-Gunn J. The association between early care arrangements, quality, and emergent bilingual Latino American children's math and literacy skills in English. *Early Childhood Research Quarterly*. 2015; 30:32–44.
- Caughy MOB, Hayslett-McCall KL, O'Campo PJ. No neighborhood is an island: incorporating distal neighborhood effects into multilevel studies of child developmental competence. *Health & Place*. 2007; 13(4):788–798. [PubMed: 17382574]
- Chertow, DS. Doctoral dissertation. 1968. Project Head Start, the urban and rural challenge: Final report submitted to Office of Economic Opportunity. Retrieved from Syracuse University
- Chetty, R.; Hendren, N.; Kline, P.; Saez, E. Where is the land of opportunity? The geography of intergenerational mobility in the United States (No. w19843). Cambridge, MA: National Bureau of Economic Research; 2014.
- Clifford RM, Barbarin O, Chang F, Early D, Bryant D, Howes C, ...Pianta R. What is pre-kindergarten? Characteristics of public pre-kindergarten programs. *Applied Developmental Science*. 2005; 9(3):126–143.
- Connors MC, Friedman-Krauss AH, Morris PA, Jones SM. From measure to construct: Refining early measures of early childhood classroom quality. 2015 Under review.

- Cooper BR, Lanza ST. Who benefits most from Head Start? Using latent class moderation to examine differential treatment effects. *Child Development*. 2014; 85(6):2317–2338. [PubMed: 25132426]
- Currie, J.; Thomas, D. Does Head Start make a difference? (No. w4406). Cambridge, MA: National Bureau of Economic Research; 1993.
- Deming D. Early childhood intervention and life-cycle skill development: Evidence from Head Start. *American Economic Journal: Applied Economics*. 2009; 1(3):111–134.
- Dickinson, DK. Large-group and free-play times: Conversational settings supporting language and literacy development?. In: Dickinson, DK.; Tabors, PO., editors. *Beginning literacy with language: Young children learning at home and at school*. Baltimore: Brookes; 2001. p. 223-255.
- Farver JAM, Lonigan CJ, Eppe S. Effective early literacy skill development for young Spanish-speaking English language learners: An experimental study of two methods. *Child Development*. 2009; 80(3):703–719. [PubMed: 19489898]
- Feller A, Grindal T, Miratrix L, Page L. Compared to what? Variation in the impacts of early childhood education by alternative care-type settings. under review.
- Garces, E.; Thomas, D.; Currie, J. Longer term effects of Head Start (No. w8054). National Bureau of Economic Research; 2000.
- Gibson, C.; Jung, K. Historical Census Statistics on the Foreign-born Population of the United States, 1850 to 2000. Population Division, US Census Bureau; 2006.
- Glaeser, E. *Triumph of the city: How our greatest invention makes us richer, smarter, greener, healthier and happier*. New York: Penguin; 2011.
- Gest SD, Holland-Coviello R, Welsh JA, Eicher-Catt DL, Gill S. Language development subcontexts in Head Start classrooms: Distinctive patterns of teacher talk during free play, meal time and book reading. *Early Education and Development*. 2006; 17(2):293–315.
- Gormley WT. The effects of Oklahoma's pre-K program on Hispanic children. *Social Science Quarterly*. 2008; 89(4):916–936.
- Grace, C.; Shores, EF.; Zaslow, M.; Brown, B.; Aufseeser, D.; Bell, L. Rural disparities in baseline data of the Early Childhood Longitudinal Study: A chartbook. (Rural Early Childhood Report No. 3). Mississippi State, MS: National Center for Rural Early Childhood Learning Initiatives, Mississippi State University Early Childhood Institute; 2006.
- Harms, T.; Clifford, M.; Cryer, D. *Early Childhood Rating Scale, Revised Edition (ECERS-R)*. New York: Teachers College Press; 1998.
- Hart, B.; Risley, TR. *Meaningful differences in the everyday experience of young American children*. Paul H Brookes Publishing; 1995.
- Hindman AH, Wasik BA. Vocabulary learning in Head Start: Nature and extent of classroom instruction and its contributions to children's learning. *Journal of School Psychology*. 2013; 51(3): 387–405. [PubMed: 23816231]
- Hubbs-Tait L, Culp AM, Huey E, Culp R, Starost HJ, Hare C. Relation of Head Start attendance to children's cognitive and social outcomes: Moderation by family risk. *Early Childhood Research Quarterly*. 2002; 17(4):539–558.
- Jencks, C.; Mayer, S. The social consequences of growing up in a poor neighborhood. In: Lynn, LE.; McGeary, MFH., editors. *Inner-city poverty in the United States*. Washington, DC: National Academy Press; 1990. p. 111-186.
- Lee K. Impacts of the duration of Head Start enrollment on children's academic outcomes: Moderation effects of family risk factors and earlier outcomes. *Journal of Community Psychology*. 2011; 39(6):698–716.
- Leventhal T, Brooks-Gunn J. The neighborhoods they live in: the effects of neighborhood residence on child and adolescent outcomes. *Psychological Bulletin*. 2000; 126(2):309. [PubMed: 10748645]
- Little, RJ.; Rubin, DB. *Statistical analysis with missing data*. New York: John Wiley & Sons; 2014.
- Lonigan CJ, Anthony JL, Bloomfield BG, Dyer SM, Samwel CS. Effects of two shared-reading interventions on emergent literacy skills of at-risk preschoolers. *Journal of Early Intervention*. 1999; 22(4):306–322.
- Magnuson K, Lahaie C, Waldfogel J. Preschool and school readiness of children of immigrants. *Social Science Quarterly*. 2006; 87(5):1241–1262.

- McKey, RH. The impact of Head Start on children, families and communities. Final report of the Head Start evaluation, synthesis and utilization project. Washington, DC: U.S. Government Printing Office; 1985.
- Miller EB, Farkas G, Vandell DL, Duncan GJ. Do the effects of Head Start vary by parental preacademic stimulation? *Child Development*. 2014; 85(4):1385–1400. [PubMed: 24597729]
- Miller P, Votruba-Drzal E. Early academic skills and childhood experiences across the urban–rural continuum. *Early Childhood Research Quarterly*. 2013; 28(2):234–248.
- National Advisory Committee on Rural Health and Human Services. Challenges to Head Start and early childhood development programs in rural communities. 2012.
- Office of Head Start. History of Head Start. Administration for Children and Families; 2014.
- Oh, SS.; Yoshikawa, H.; Willett, JB. Unpacking the black box of Head Start: Investigating the mediational processes of parental investments of program impact on children's language, literacy & math outcomes. Paper presented at the Society for Research in Child Development (SRCD) Biennial Meeting; Philadelphia, PA. 2015 Mar.
- Pinderhughes EE, Nix R, Foster EM, Jones D. The Conduct Problems Prevention Research Group. Parenting in context: Impact of neighborhood poverty, residential stability, public services, social networks, and danger on parental behaviors. *Journal of Marriage and Family*. 2001; 63(4):941–953. [PubMed: 19829752]
- Puma, M.; Bell, S.; Cook, R.; Heid, C.; Shapiro, G.; Broene, P.; Spier, E. Head Start Impact Study. Final Report. Washington, DC: Administration for Children & Families; 2010a.
- Puma, M.; Bell, S.; Cook, R.; Heid, C.; Shapiro, G.; Broene, P.; Spier, E. Head Start Impact Study. Technical Report. Washington, DC: Administration for Children & Families; 2010b.
- Raudenbush SW, Reardon SF, Nomi T. Statistical analysis for multisite trials using instrumental variables with random coefficients. *Journal of Research on Educational Effectiveness*. 2012; 5(3): 303–332.
- Raudenbush, S.; Bloom, H. Learning about and from variation in program impacts using multisite trials. New York: MDRC; 2015.
- Resnick, G.; Zill, N. Is Head Start providing high-quality educational services? Unpacking classroom processes. Westat, Inc; 2000. undated xerox
- Robinson JL, Emde RN. Mental health moderators of Early Head Start on parenting and child development: Maternal depression and relationship attitudes. *Parenting: Science and Practice*. 2004; 4(1):73–97.
- Rural Poverty Research Institute. Human service programs serving rural communities: Head Start. Corvallis, OR: Rural Poverty Research Institute; 2008.
- Rutter M. The city and the child. *American Journal of Orthopsychiatry*. 1981; 51(4):610. [PubMed: 7294164]
- Sampson RJ, Morenoff JD, Earls F. Beyond social capital: Spatial dynamics of collective efficacy for children. *American Sociological Review*. 1999:633–660.
- Sampson RJ, Morenoff JD, Gannon-Rowley T. Assessing “neighborhood effects”: Social processes and new directions in research. *Annual Review of Sociology*. 2002:443–478.
- Shager HM, Schindler HS, Magnuson KA, Duncan GJ, Yoshikawa H, Hart CM. Can research design explain variation in Head Start research results? A meta-analysis of cognitive and achievement outcomes. *Educational Evaluation and Policy Analysis*. 2013; 35(1):76–95.
- Shinn M, Toohey SM. Community contexts of human welfare. *Annual Review of Psychology*. 2003; 54(1):427–459.
- Smith, K. Rural families choose home-based child care for their preschool children. Durham, NH: Carsey Institute; 2006.
- Smith MS, Bissell JS. Report analysis: The impact of Head Start. *Harvard Educational Review*. 1970; 40(1):51–104.
- Spybrook J. Detecting intervention effects across context: An examination of the precision of cluster randomized trials. *The Journal of Experimental Education*. 2014; 82(3):334–357.
- Swenson, K. Child care arrangements in urban and rural areas. Washington, DC: US Department of Health and Human Services; 2008.

- Teale WH, Paciga KA, Hoffman JL. Issues in urban literacy: Beginning reading instruction in urban schools: The curriculum gap ensures a continuing achievement gap. *The Reading Teacher*. 2007;344–348.
- Tieken, MC. *Why Rural Schools Matter*. UNC Press Books; 2014.
- Von Hippel PT. Regression with missing Ys: An improved strategy for analyzing multiply imputed data. *Sociological Methodology*. 2007; 37(1):83–117.
- Zhang Y. Urban-rural literacy gaps in Sub-Saharan Africa: The roles of socioeconomic status and school quality. *Comparative Education Review*. 2006; 50(4):581–602.

Research Highlights

- We examine variation in Head Start's impact on language and literacy by urbanicity
- Head Start is more effective in improving receptive vocabulary in urban areas
- Head Start improves children's oral comprehension only in rural areas
- Related contextual and compositional features may underlie urban-rural differences

Figure 1.

Impacts of random assignment to Head Start on receptive vocabulary scores in urban, mixed, and rural communities

Notes: ** $p < .01$, * $p < .05$, + $p < .10$, ns $p \geq .10$; Chars = Characteristics; Significance levels in parentheses next to axis labels indicate significance of urbanicity effect on impact of Head Start random assignment; Numbers and significance levels of labels on each bar indicate effect size and significance level of the impact of Head Start in a given context and model; Urban neighborhoods defined as those with 100% of households falling in urban area/cluster, mixed neighborhoods defined as those with 50% of households in urban and 50% in rural area/cluster, and rural neighborhoods defined as those with 0% of households in urban area/cluster

Figure 2.

Impacts of random assignment to Head Start on oral comprehension scores in urban and rural communities

Notes: * $p < .05$, + $p < .10$, ns $p \geq .10$; Chars = Characteristics; Significance levels in parentheses next to axis labels indicate significance of urbanicity effect on impact of Head Start random assignment; Numbers and significance levels of labels on each bar indicate effect size and significance level of the impact of Head Start in a given context and model; Urban neighborhoods defined as those with 100% of households falling in urban area/cluster, mixed neighborhoods defined as those with 50% of households in urban and 50% in rural area/cluster, and rural neighborhoods defined as those with 0% of households in urban area/cluster

Table 1

Descriptive statistics and correlations with urbanicity

	Sample N before imputation	% missing before imputation	Mean/%	Std Dev	Min	Max	r with urbanicity
<i>Child and family characteristics^a</i>							
Spring PPVT	3503	0%	272.89	40.09	154.07	401.42	-.10**
Spring WJ Oral Comprehension	3503	0%	439.04	16.38	418.00	489.00	-.18**
Spring WJ Letter Word Identification	3503	0%	313.26	27.80	264.00	408.00	.02
Spring WJ Spelling	3503	0%	357.23	26.03	277.00	442.00	.06**
Treatment (Head Start = 1)	3503	0%	63%				.00
Fall PPVT	3095	12%	250.45	42.38	128.54	388.11	-.06**
Fall WJ Oral Comprehension	2318	34%	438.62	14.65	418.00	489.00	-.06**
Fall WJ Letter Word Identification	3064	13%	300.73	24.71	264.00	396.00	.05**
Fall WJ Spelling	2366	32%	344.49	27.00	277.00	432.00	.04*
Child age (months)	3503	0%	4.04	0.66	2.00	6.00	.07**
Child age cohort (1 = 3-year-old)	3503	0%	55%				-.06**
Child male	3503	0%	50%				.00
Child black	3487	0%	31%				-.01
Child Hispanic	3487	0%	35%				.25**
English as home language	3475	1%	72%				-.28**
Single mother	3378	4%	55%				-.06**
Mother less than HS education	3375	4%	37%				-.00
Mother recent immigrant	3464	1%	17%				.09**
Maternal depressive symptoms	1135	68%	1.78	0.98	1.00	4.00	.21**
<i>Center characteristics^b</i>							
Proportion teachers with BA	274	14%	0.15	0.25	0.00	1.00	.21**
Transportation provided	277	13%	60%				-.30**
Vacant teaching slot(s)	277	13%	6%				.05
Full-day care provided	277	13%	62%				.09+

	Sample N before imputation	% missing before imputation	Mean%	Std Dev	Min	Max	r with urbanicity
Family and child services	277	13%	0.65	0.22	0.05	0.95	-.13*
Mean classroom quality (materials & space)	294	7%	0.67	0.14	0.13	0.99	-.02
Mean classroom quality (positive interactions)	294	7%	0.77	0.15	0.22	1.00	-.07
Mean classroom quality (negative interactions)	294	7%	0.05	0.09	0.00	0.59	.01
Mean classroom literacy activities	301	5%	0.64	0.22	0.00	1.00	-.01
<i>Center neighborhood characteristics:^b</i>							
Urbanicity (proportion)	317	0%	0.66	0.46	0.00	1.00	--
Poverty (proportion)	315	1%	0.24	0.14	0.02	0.79	.24**
Crimes / 1000 people	314	1%	42.51	17.82	2.75	81.19	.45**
Ethnic minority concentration (proportion)	315	1%	0.41	0.32	0.01	1.00	.44**
DOE preschools / 100 3–5 year olds	315	1%	0.12	0.14	0.00	0.77	-.02
Total resources	315	1%	261.81	207.43	2.00	1113.00	.48**

Notes:

**
 $p < .01$,*
 $p < .05$,+
 $p < .10$;^a Measured at individual level, with total possible $n = 3503$;^b Measured at random assignment center level, with total possible $n = 317$

Table 2

Two-level regression models showing the differential impact of Head Start random assignment on child receptive vocabulary scores based on center neighborhood urbanicity

	Basic Model			With Neighborhood Characteristics			With Center Characteristics			With Compositional Characteristics		
	B	SE	p	B	SE	p	B	SE	p	B	SE	p
<i>Level 1 predictors</i>												
Treatment assignment	5.34	0.97	**	5.30	0.96	**	5.28	0.98	**	5.28	0.95	**
Pretest score	0.45	0.01	**	0.45	0.01	**	0.45	0.01	**	0.46	0.02	**
Child age	7.18	0.92	**	7.19	0.92	**	7.15	0.92	**	11.12	1.57	**
Child age cohort	-13.66	1.25	**	-13.58	1.25	**	-13.64	1.25	**	-11.17	2.11	**
Child male	-2.20	0.90	**	-2.28	0.90	**	-2.26	0.90	**	-3.72	1.57	+
Child black	-4.07	1.94	*	-3.96	1.94	*	-4.14	1.95	*	-0.86	2.61	
Child Hispanic	-1.36	1.81		-1.38	1.81		-1.35	1.81		-3.81	2.66	
Home language (1 = English)	16.64	1.86	**	16.67	1.86	**	16.61	1.86	**	15.97	2.98	**
Single mother	-1.62	0.96	+	-1.55	0.96		-1.64	0.96	+	-2.10	1.67	
Mother less than HS education	-5.93	1.00	**	-6.03	1.00	**	-5.92	1.00	**	-5.99	1.72	**
Mother recent immigrant	-3.85	1.62	*	-3.89	1.62	*	-3.98	1.62	*	-6.84	2.86	*
Maternal depressive symptoms	0.49	0.46		0.50	0.46		0.48	0.46		0.25	0.64	
<i>Level 2 predictors of treatment assignment</i>												
<i>Neighborhood</i>												
Urbanicity	3.64	2.15	+	4.01	2.79		3.36	2.33		2.38	2.24	
Poverty				-16.12	8.63	+						
Crime				-0.13	0.07	*						
Ethnic minority concentration				6.90	4.29							
DOE preschools				-2.54	6.50							
Total resources				0.01	0.01							
<i>Center</i>												
% teachers with BA							-4.01	3.94				
Transportation provided							-1.68	2.19				
Vacant teaching slots							4.06	3.64				
Ease of replacing teachers							-0.29	1.07				

	Basic Model			With Neighborhood Characteristics			With Center Characteristics			With Compositional Characteristics		
	B	SE	p	B	SE	p	B	SE	p	B	SE	p
Full-day care provided							-1.21	2.13				
Family and child services							-4.52	4.50				
Mean material & spatial quality							8.79	9.89				
Mean positive interactions							-2.10	9.81				
Mean negative interactions							-0.07	15.01				
Mean literacy activities							-5.18	4.99				
<i>Level 1 predictors interacted with treatment assignment</i>												
Treatment * Pretest score										-0.02	0.03	
Treatment * Child age										-5.74	1.87	**
Treatment * Child age cohort										-3.37	2.49	
Treatment * Child male										1.98	1.89	
Treatment * Child black										-4.51	2.55	+
Treatment * Child Hispanic										3.59	2.93	
Treatment * Home language										1.17	3.49	
Treatment * Single mother										0.82	2.01	
Treatment * Mother education										0.31	2.07	
Treatment * Mother immigrant										4.72	3.38	
Treatment * Maternal depression										0.39	0.64	

Notes:

**
 $p < .01$,*
 $p < .05$,+
 $p < .10$; Bold text for urbanicity indicates primary variable of interest

Table 3

Two-level regression models showing the differential impact of Head Start random assignment on child oral comprehension scores based on center neighborhood urbanicity

	Basic Model			With Neighborhood Characteristics			With Center Characteristics			With Compositional Characteristics		
	B	SE	p	B	SE	p	B	SE	p	B	SE	p
<i>Level 1 predictors</i>												
Treatment assignment	0.29	0.42		0.29	0.42		0.31	0.43		0.32	0.43	
Pretest score	0.40	0.02	**	0.40	0.02	**	0.40	0.02	**	0.41	0.02	**
Child age	2.68	0.40	**	2.70	0.40	**	2.67	0.40	**	2.37	0.68	**
Child age cohort	-2.29	0.55	**	-2.26	0.55	**	-2.30	0.55	**	-2.62	0.91	**
Child male	-1.11	0.41	**	-1.12	0.41	**	-1.10	0.41	**	-1.38	0.69	*
Child black	-0.33	0.87		-0.29	0.87		-0.35	0.88		-0.06	1.16	
Child Hispanic	-1.66	0.81	*	-1.70	0.81	*	-1.67	0.81	*	-1.90	1.18	
Home language (1 = English)	11.37	0.83	**	11.37	0.83	**	11.39	0.83	**	10.74	1.31	**
Single mother	0.25	0.43		0.26	0.43		0.25	0.43		0.45	0.74	
Mother less than HS education	-2.07	0.45	**	-2.101	0.45	**	-2.07	0.45	**	-2.65	0.76	**
Mother recent immigrant	-1.50	0.73	*	-1.49	0.73	*	-1.47	0.73	*	-1.56	1.27	
Maternal depressive symptoms	-0.02	0.21		-0.02	0.21		-0.02	0.21		-0.32	0.28	
<i>Level 2 predictors of treatment assignment</i>												
<i>Neighborhood</i>												
Urbanicity	-2.03	0.94	*	-1.81	1.23		-1.37	1.02		-1.94	1.01	*
Poverty				-6.46	3.81	+						
Crime				-0.03	0.03							
Ethnic minority concentration				1.07	1.89							
DOE preschools				2.64	2.88							
Total resources				0.00	0.00							
<i>Center</i>												
% teachers with BA							-1.44	1.72				
Transportation provided							1.74	0.96	+			
Vacant teaching slots							-0.49	1.59				
Ease of replacing teachers							0.11	0.47				

	Basic Model			With Neighborhood Characteristics			With Center Characteristics			With Compositional Characteristics		
	B	SE	p	B	SE	p	B	SE	p	B	SE	p
Full-day care provided							0.71	0.93				
Family and child services							-0.84	1.96				
Mean material & spatial quality							-1.35	4.31				
Mean positive interactions							3.70	4.29				
Mean negative interactions							2.86	6.56				
Mean literacy activities							1.08	2.18				
<i>Level 1 predictors interacted with treatment assignment</i>												
Treatment * Pretest score										-0.02	0.03	
Treatment * Child age										0.46	0.82	
Treatment * Child age cohort										0.49	1.08	
Treatment * Child male										0.40	0.85	
Treatment * Child black										-0.41	1.13	
Treatment * Child Hispanic										0.37	1.31	
Treatment * Home language										0.92	1.54	
Treatment * Single mother										-0.31	0.90	
Treatment * Mother education										0.85	0.92	
Treatment * Mother immigrant										0.09	1.51	
Treatment * Maternal depression										0.44	0.29	

Notes:

**
 $p < .01$,*
 $p < .05$,+
 $p < .10$; Bold text for urbanicity indicates primary variable of interest

Two-level regression models showing the differential impact of Head Start random assignment on child pre-reading scores based on center neighborhood urbanicity

Table 4

	Basic Model			With Neighborhood Characteristics			With Center Characteristics			With Compositional Characteristics		
	B	SE	p	B	SE	p	B	SE	p	B	SE	p
<i>Level 1 predictors</i>												
Treatment assignment	4.58	0.86	**	4.59	0.86	**	4.64	0.86	**	4.60	0.86	**
Pretest score	0.54	0.02	**	0.54	0.02	**	0.54	0.02	**	0.55	0.03	**
Child age	4.55	0.77	**	4.51	0.77	**	4.55	0.77	**	3.01	1.27	*
Child age cohort	-5.98	1.05	**	-6.02	1.05	**	-6.01	1.05	**	-7.85	1.67	**
Child male	-3.35	0.79	**	-3.36	0.79	**	-3.37	0.79	**	-2.54	1.29	*
Child black	0.16	1.68		0.13	1.68		0.26	1.69		-1.69	2.19	
Child Hispanic	-1.86	1.56		-1.79	1.56		-1.95	1.56		-2.56	2.19	
Home language (1 = English)	2.22	1.59		2.23	1.60		2.14	1.59		2.62	2.43	
Single mother	-1.01	0.83		-0.99	0.83		-1.00	0.83		-0.69	1.36	
Mother less than HS education	-3.24	0.87	**	-3.21	0.87	**	-3.23	0.87	**	-5.26	1.41	**
Mother recent immigrant	0.76	1.41		0.72	1.41		0.71	1.41		0.49	2.34	
Maternal depressive symptoms	0.37	0.40		0.38	0.40		0.41	0.40		0.21	0.54	
<i>Level 2 predictors of treatment assignment</i>												
<i>Neighborhood</i>												
Urbanicity	-1.90	1.90		-1.89	2.48		-1.90	2.03		2.82	2.03	
Poverty				7.01	7.75							
Crime				-0.02	0.06							
Ethnic minority concentration				1.42	3.83							
DOE preschools				-5.05	5.78							
Total resources				-0.00	0.01							
<i>Center</i>												
% teachers with BA							4.51	3.43				
Transportation provided							1.55	1.91				
Vacant teaching slots							3.58	3.17				
Ease of replacing teachers							0.84	0.93				

	Basic Model			With Neighborhood Characteristics			With Center Characteristics			With Compositional Characteristics		
	B	SE	p	B	SE	p	B	SE	p	B	SE	p
Full-day care provided							-0.38	1.86				
Family and child services							-2.59	3.93				
Mean material & spatial quality							3.18	8.61				
Mean positive interactions							5.69	8.53				
Mean negative interactions							21.54	13.08	+			
Mean literacy activities							4.70	4.33				
<i>Level 1 predictors interacted with treatment assignment</i>												
Treatment * Pretest score										-0.00	0.03	
Treatment * Child age										2.32	1.56	
Treatment * Child age cohort										2.91	2.05	
Treatment * Child male										-1.24	1.61	
Treatment * Child black										2.72	2.20	
Treatment * Child Hispanic										1.10	2.49	
Treatment * Home language										-0.74	2.93	
Treatment * Single mother										-0.51	1.70	
Treatment * Mother education										3.17	1.74	+
Treatment * Mother immigrant										0.42	2.86	
Treatment * Maternal depression										0.22	0.56	

Notes:

**
 $p < .01$,*
 $p < .05$,+
 $p < .10$; Bold text for urbanicity indicates primary variable of interest

Table 5

Two-level regression models showing the differential impact of Head Start random assignment on child pre-writing scores based on center neighborhood urbanicity

	Basic Model			With Neighborhood Characteristics			With Center Characteristics			With Compositional Characteristics		
	B	SE	p	B	SE	p	B	SE	p	B	SE	p
<i>Level 1 predictors</i>												
Treatment assignment	2.72	0.79	**	2.70	0.79	**	2.73	0.80	**	6.75	9.36	
Pretest score	0.36	0.02	**	0.36	0.02	**	0.36	0.02	**	0.37	0.02	**
Child age	6.06	0.75	**	6.05	0.75	**	6.04	0.75	**	6.73	1.23	**
Child age cohort	-9.13	1.00	**	-9.12	1.00	**	-9.13	1.00	**	-9.16	1.62	**
Child male	-4.90	0.75	**	-4.93	0.75	**	-4.89	0.75	**	-4.21	1.25	**
Child black	-2.58	1.59		-2.58	1.59		-2.56	1.60	+	-1.43	2.09	
Child Hispanic	-1.64	1.48		-1.61	1.48		-1.67	1.48		-1.34	2.11	
Home language (1 = English)	-0.58	1.51		-0.54	1.51		-0.55	1.51		0.94	2.35	
Single mother	-1.76	0.79	*	-1.78	0.79	*	-1.79	0.79	*	-2.59	1.31	*
Mother less than HS education	-2.35	0.82	**	-2.37	0.82	**	-2.35	0.82	**	-3.67	1.36	**
Mother recent immigrant	-0.50	1.33		-0.53	1.33		-0.48	1.34		-0.02	2.26	
Maternal depressive symptoms	-0.37	0.38		-0.36	0.38		-0.37	0.38		-0.48	0.51	
<i>Level 2 predictors of treatment assignment</i>												
<i>Neighborhood</i>												
Urbanicity	0.44	1.74		1.44	2.30		1.10	1.90		0.06	1.87	
Poverty				-0.22	7.16							
Crime				-0.01	0.06							
Ethnic minority concentration				-1.45	3.54							
DOE preschools				-5.60	5.37							
Total resources				-0.00	0.00							
<i>Center</i>												
% teachers with BA							-0.58	3.21				
Transportation provided							2.25	1.79				
Vacant teaching slots							-1.68	2.97				
Ease of replacing teachers							-0.46	0.87				

	Basic Model			With Neighborhood Characteristics			With Center Characteristics			With Compositional Characteristics		
	B	SE	p	B	SE	p	B	SE	p	B	SE	p
Full-day care provided							0.06	1.74				
Family and child services							-2.27	3.68				
Mean material & spatial quality							-1.86	8.07				
Mean positive interactions							4.81	8.00				
Mean negative interactions							4.34	12.25				
Mean literacy activities							1.94	4.07				
<i>Level 1 predictors interacted with treatment assignment</i>												
Treatment * Pretest score										-0.01	0.03	
Treatment * Child age										-1.07	1.50	
Treatment * Child age cohort										0.10	1.96	
Treatment * Child male										-1.08	1.54	
Treatment * Child black										-1.71	2.06	
Treatment * Child Hispanic										-0.50	2.36	
Treatment * Home language										-2.36	2.80	
Treatment * Single mother										1.32	1.62	
Treatment * Mother education										2.09	1.66	
Treatment * Mother immigrant										-0.58	2.73	
Treatment * Maternal depression										0.19	0.53	

Notes:

**
 $p < .01$,*
 $p < .05$,+
 $p < .10$; Bold text for urbanicity indicates primary variable of interest