American Educational Research Journal
Month XXXX, Vol. XX, No. X, pp. 1–35
DOI: 10.3102/0002831219872977
Article reuse guidelines: sagepub.com/journals-permissions

The Moderating Effect of Neighborhood Poverty on Preschool Effectiveness: Evidence From the Tennessee Voluntary Prekindergarten Experiment

Francis A. Pearman II
Stanford University

This study drew data from a randomized trial of a statewide prekindergarten program in Tennessee and presents new evidence on the impacts of preK on third-grade achievement using administrative data on children's neighborhood environments. Results indicate that preK had no measurable impact on children's third-grade math achievement regardless of children's neighborhood conditions. However, preK significantly improved third-grade reading achievement for children living in high-poverty neighborhoods. The treatment effects on reading achievement were substantial: Among children living in high-poverty neighborhoods, those who took up an experimental assignment to attend preK scored over half a standard deviation higher on average than the control group in third grade. In contrast, preK enrollment had, if anything, a negative effect on third-grade reading achievement among children living in low-poverty neighborhoods. These differential effects were partially explained by alternative childcare options and contextual risk factors.

KEYWORDS: academic achievement, moderation analysis, neighborhood poverty, preK, randomized controlled trial

Introduction

Experimental evidence for the positive long-term effects of preschool comes from a small set of interventions implemented in confined geographical areas. For instance, the Perry Preschool Project, commonly cited as a model of effective preK, served low-income children living in the attendance zone of a single high-poverty elementary school in Ypsilanti, MI

FRANCIS A. PEARMAN II is an assistant professor of education at Stanford University, CERAS 524, 485 Lasuen Mall, Stanford, CA 94305, USA; e-mail: *apearman@stanford.edu*. His research focuses on social and educational inequality.

(Berruate-Clement, 1984). In contrast, states looking to expand access to prekindergarten for large segments of their child population will need to accommodate children living in neighborhoods that differ markedly from those served by the Perry Preschool Project. Despite an extensive neighborhood effects literature that has documented qualitative differences in the experience of poverty across places (Small, 2008; Votruba-Drzal, Miller, & Coley, 2016), researchers have largely assumed away what these contextual differences might mean for preschool effects.

This oversight is notable because it is reasonable to suspect that variation in neighborhood conditions may have implications for what children experience in the absence of preK—what is commonly referred to as the *counterfactual condition* (Bloom & Weiland, 2015). For instance, if a low-income child grows up in a high-poverty neighborhood relative to a more affluent or socioeconomically mixed neighborhood, existing evidence suggests that this child will likely experience heightened stress (Brooks-Gunn, Johnson, & Leventhal, 2010), increased exposure to violence and victimization (Sharkey, Tirado-Strayer, Papachristos, & Raver, 2012), greater likelihood of parental disengagement (Cohen, 2017), as well as fewer quality day care and out-of-school learning opportunities (Jacob & Wilder, 2010). Should there exist systematic variation in what children experience in the absence of preK that is tied to the structure and function of residential contexts, then even a standardized preK program implemented with fidelity across a broad range of locations may still produce variation in treatment effects across neighborhood contexts. Despite the potential importance of understanding this brand of impact heterogeneity in an era of preK expansion (Morris et al., 2017)—in particular, for informing decisions about how best to target preK interventions—little is known as to whether residential environments matter for how low-income children engage and respond to preschool.

This article revisits data from the Tennessee Voluntary Prekindergarten Program (TN-VPK), the first randomized trial of a scaled-up, statewide prekindergarten initiative in the United States. Data for this study were drawn from the full sample of eligible, randomly assigned children, as well as from a smaller subsample for whom a broader set of baseline measures was available. This study leverages the fact that there existed large variation in the level of poverty in children's residential neighborhoods at baseline, even though the sample itself was composed exclusively of economically disadvantaged children. This variation made it possible to explore whether the effect of preschool on later achievement varied based on where children lived. In particular, the following research question guides this study: Do the effects of assignment to or enrollment in preschool on children's third-grade achievement vary across levels of poverty in children's residential neighborhoods at baseline? Additionally, this study attempts to shed light on factors that might be responsible for potential impact heterogeneity across neighborhood contexts, including differential access to alternative childcare options, differential exposure to contextual risk factors, and differences in family socioeconomic status (SES) across neighborhood contexts.

This study advances research on preK education by providing the first estimates, to the author's knowledge, of whether preK effects on later academic achievement differ based on the characteristics of children's residential neighborhoods. This article begins with a brief review of why neighborhood poverty matters for children's academic achievement. Next, this paper describes relevant early childhood literature examining impact heterogeneity across community characteristics. After providing a brief overview of the Tennessee Voluntary PreK program, the paper turns attention to the data, randomization method, and the empirical strategy used to answer the question of whether preK effects differ across levels of neighborhood poverty. Finally, results are presented and possible explanations are explored before discussing implications for policy and research.

Background

A considerable body of research has underscored the notion that neighborhoods play a prominent role in influencing children's early development (Chetty, Hendren, & Katz, 2015; Wodtke, Elwert, & Harding, 2016). However, surprisingly little attention has been devoted to the role that neighborhood environments play in shaping who benefits from early childhood education. As noted previously, this literature gap is notable because the opportunities and risk factors embedded in neighborhood environments help shape the nature of the counterfactual condition to which preschool exposure is compared.

Neighborhoods, Preschool, and Academic Achievement

There are at least three reasons why neighborhoods may influence the extent to which children benefit from a statewide preschool program. First, neighborhoods differ in alternative childcare options (Swenson, 2008). Children in high-poverty neighborhoods are less likely than their peers in more affluent neighborhoods to have access to nursery schools, private or informal day care facilities, or other enrichment programs (Green, 2015; Jocson & Thorne-Wallington, 2013; Tate, 2008). Moreover, there is evidence that alternative childcare options in high-poverty neighborhoods are of lower quality than those in more affluent neighborhoods (McCoy, Connors, Morris, Yoshikawa, & Friedman-Krauss, 2015; Valentino, 2018). Therefore, it is conceivable that the effects of a statewide preK program could differ depending on whether children have limited access to community-based alternatives or if their access is circumscribed by lower quality services, with preK effects potentially being most pronounced if so.

Second, neighborhoods can differ in terms of contextual risk factors, including exposure to violence, housing instability, unemployment, and weakened family units (Brooks-Gunn et al., 2010; Sharkey et al., 2012). If preK effects operate, in part, by safeguarding children from developmental risk factors during periods of elevated developmental vulnerability (Dodge, Greenberg, & Malone, 2008), then one would expect preK to be most beneficial for those in high-risk communities where exposure to preK can presumably do the most safeguarding. Finally, prior research has indicated that preK effects are largest for children from economically disadvantaged families (Cooper & Lanza, 2014; Gormley, Phillips, & Gayer, 2008; Yoshikawa et al., 2013). Therefore, given that disadvantaged families often live in disadvantaged neighborhoods (Reardon & Bischoff, 2011), any independent, moderating effect of family SES on preschool effects would be observed to some extent at the neighborhood level.

The few studies that have directly tested whether preK effects hinge on neighborhood characteristics have used broad distinctions between urban and rural communities and have operationalized community characteristics based on the location of the preK center itself. Fitzpatrick (2008) drew observational data from the National Assessment of Educational Progress and found, based on a regression–discontinuity design, that disadvantaged children attending preK in small towns and rural areas of Georgia experienced the largest gains in math and reading scores after the expansion of universal preK. By contrast, McCoy, Morris, Connors, Gomez, and Yoshikawa (2016) found that Head Start was more effective at improving receptive vocabulary scores in urban areas but was more effective at improving oral comprehension scores in rural areas. Although these studies provide important insights into preK effect moderation by community characteristics in a general sense, their precise insights into how neighborhood conditions relate to preK effects are limited for two reasons.

First, although the urban-rural divide may have implications for the types of risk and protective factors embedded in each community type (Votruba-Drzal et al., 2016), this binary conceals important heterogeneity that exists within each category—heterogeneity that captures more directly the level of disadvantage that children encounter in their residential areas. After all, the term "urban" spans not only central city poverty but also *sub*-urban wealth (Milner, 2015), while the term "rural" similarly encompasses a broad range of communities that differ in socioeconomic composition (Flora, Flora, & Gasteyer, 2015). Second, prior preschool research on impact heterogeneity by community characteristics has focused on the location of the preK center itself rather children's own residential neighborhoods. This is notable because measuring disadvantage at the level of children's residential environment is likely a better approximation of the contextual determinants of children's counterfactual condition than the area in which the preK center is located, which may or may not be near children's homes.

In fact, the neighborhood environment of the preK center could differ markedly from children's home neighborhood and thus obscure what is understood about the counterfactual condition.

For instance, as shown in the Supplementary Appendix Table A.1 in the online version of the journal, approximately 44% of children in the full analytic sample who lived in high-poverty neighborhoods attempted to enroll (and were subsequently randomized) in preK centers that were located in low-poverty neighborhoods. (A full description of the data and randomization method is described below.) Likewise, 7% of children living in low-poverty neighborhoods were randomized at preK centers located in high-poverty neighborhoods. These crossovers were permissible because enrollment in the TN-VPK program, unlike many K–12 schools in Tennessee, was not based on catchment area assignment. Consequently, many children attempted to enroll in preK centers that were not necessarily local, reiterating the importance of capturing counterfactual variation in terms of residential contexts.

To assess variation in neighborhood conditions and examine whether this variation moderates the effect of preK, this study revisits data from a randomized trial of a statewide prekindergarten initiative, the TN-VPK program. Recent research on TN-VPK found that initial gains from preK faded by third grade (Lipsey, Farran, & Durkin, 2017), a finding echoed in several recent experimental studies of preschool effects (Puma, Bell, Cook, & Heid, 2010; Puma et al., 2012). The current study reexamines TN-VPK findings with an eye toward whether residential contexts mattered in these achievement patterns. In particular, this study asks whether neighborhood poverty moderates the effect of preK on children's achievement in third grade.

Tennessee Voluntary PreK Program

The Tennessee Voluntary Prekindergarten (TN-VPK) program, a statewide early childhood initiative housed in the Tennessee Department of Education (TNDOE), enrolls over 18,000 4-year-olds in school districts across Tennessee. The structure of TN-VPK is similar to other statewide prekindergarten programs implemented within state departments of education (Karch, 2010). In particular, local school districts submit applications to TNDOE for funding to establish preK classrooms. TNDOE awards local school districts a funding amount based on the state's *Basic Education Program* formula, which accounts for a number of factors including a district's financial need. Districts are then expected to provide matching funds in order to cover the full cost of establishing the prekindergarten program. However, districts are permitted to allocate additional dollars if they wish to provide supplementary services such as transportation that are not required by TNDOE.

TN-VPK serves 4-year-old children eligible for kindergarten the following year whose families meet a number of eligibility requirements. Children eligible for the free-and-reduced-price lunch program are given top priority for admission. As space permits, children are also eligible if they are English language learners, in state custody, deemed unserved or underserved by a state advisory council, or have a disability. Per state mandates, TN-VPK requires 5 days of teaching with a minimum of 5.5 hours of instruction per day. Maximum class size is 20 students, with one teacher required for every 10 students. All VPK classrooms are required to be staffed by at least one teacher who is state-licensed and endorsed for Early Child Education by the state board of education. Additionally, preference is for at least one assistant per classroom to hold a child development certificate or an associate degree in early childhood.

Each classroom is required to operate a comprehensive curriculum that is approved by the Tennessee Department of Education's Office of Early Learning. These curricula must incorporate instruction in a broad range of developmental domains, including linguistic, cognitive, social-personal, and physical, and must be delivered through a mix of direct and individualized instruction, group activities, and center-based activities. The list of approved curricula during the 2014-2015 school year included 22 "comprehensive" curricula and 12 supplementary curricula. To ensure quality implementation of the curricula, TN-VPK drew from standards set forth by the National Institute of Early Education Research and are described in detail in the TNDOE's Scope of Services (TNDOE, 2017). TN-VPK was determined to have met 9 of the 10 current standards advocated by the National Institute of Early Education Research. These standards and curricular mandates ensure a minimum level of quality across TN-VPK classrooms.[1]

Method

This study is situated in a larger evaluation of the TN-VPK program. The evaluation began during the 2009–2010 school year and leveraged the fact that many preschools across the state were oversubscribed. This oversubscription meant that many eligible children were denied admission. At the start of the 2009–2010 school year and again at the start of the subsequent school year, a group of sites identified as oversubscribed by TNDOE agreed to randomly assign applicants to a treatment condition that was granted admission and a control condition that was denied admission. This randomization produced across two cohorts of children during the 2009–2010 and 2010–2011 school years a total of 111 randomized applicant lists in 79 schools in 29 school districts.

Of the 3,131 eligible children in these randomized applicant lists, 2,990 were included in the state education database for at least 1 school year after prekindergarten. There is no evidence that the remaining children attended public schools in Tennessee; however, the parents of 11 of these children who were either home-schooled or enrolled in private school provided consent to be followed in the Intensive Substudy (ISS) sample (described in more detail below).

These 2,990 children composed the *full randomized sample* (hereafter referred to as "the RCT sample"). The randomized controlled trial (RCT) sample includes all children who were eligible for preK enrollment at the beginning of their respective prekindergarten year and who were randomly assigned to attend or not to attend VPK. In all, 1,852 children received offers of admission, and the remaining 1,138 did not. These two groups made up the intent-to-treat (ITT) treatment and control conditions for the RCT analysis. In addition, state databases were used to identify which of these students actually enrolled in VPK, regardless of their assignment status. Of the 2,990 children in the sample, 1,997 attended VPK for at least 1 day; the remaining 993 children had no VPK attendance during their respective prekindergarten year.

Intensive Substudy Sample

The ISS sample is a subset of the full sample for whom parental consent was obtained to collect annual assessment data through third grade as well as a broader set of baseline characteristics, including achievement scores and indicators of family SES. Attempts were made to contact parents of every child on eligible randomization lists at the start of their respective prekindergarten year, and few parents explicitly refused. However, logistical challenges limited the consent rate for the first cohort of children. In particular, for the 2009–2010 school year, consent could only be obtained by way of mail-in responses that were mailed to parents directly from TNDOE before the start of the school year. As a result, the consent rate for the first cohort was 24.4%.

For the second cohort, arrangements were made to obtain parental consent during the VPK application process, which was in-person and did not require any additional steps. This modification led to a higher consent rate of 67.9% for the second cohort. These procedures together yielded 1,331 consented children across both cohorts. However, additional restrictions including limiting the sample to age- and income-eligible children and restricting children to those on randomization lists that contained at least one consented child in the treatment and control condition yielded a final analytic sample of 1,076 children, who came to be known as the "Intensive Substudy" (ISS) sample. Of these children, 697 were randomly assigned to VPK, and the remaining 379 to the control group. (Supplementary Appendix Table E.1 shows that compliance rates for the RCT and ISS samples did not differ across levels of neighborhood poverty.)

As described by Lipsey, Hofer, Dong, Farran, and Bilbrey (2013), consent rates differed by treatment condition across both cohorts. In particular, parents of children randomly assigned to VPK were more likely to consent into the ISS sample than parents of children who were not assigned to VPK. (Consent rates for Cohorts 1 and 2 differed by 14 and 6 percentage points between treatment conditions, respectively.) The differential consent rates across treatment conditions for both cohorts introduced potential bias into the ISS sample with respect to the treatment-control contrast, despite that this subsample was drawn from a randomized applicant list. As described below, the analytic models used in the current study control for a host of baseline characteristics in an effort to mitigate potential bias within this subsample. (Exploratory analysis, in the Supplementary Appendix Table H.1, revealed no evidence that treatment-control differences in consent rates differed across levels of neighborhood poverty.) Moreover, to improve generalizability of the ISS sample with regard to the full RCT sample, consent weights were computed that effectively upweighted children least likely to consent into the ISS sample. These weights were computed from a logistic regression in which a binary indicator for consenting into the ISS sample was regressed on a full set of baseline characteristics. (Complete regression results are provided in the Supplementary Appendix Table F.1.) This regression was used to compute predicted probabilities of consent for each child. The inverse of these predicted probabilities became the consent weights, which were used in subsequent analyses for the ISS sample.

Geolocating Residential Addresses

On children's initial application to TN-VPK, parents were required to indicate a home address, which was subsequently geocoded and matched to block-level data. However, not every parent provided a valid street address that could be geocoded. Of the 2,990 children in the RCT sample, 86% could be geocoded; of the 1,076 children in the ISS sample, 85% could be geocoded. (Supplementary Appendix Table G.1 shows no evidence that the likelihood of having a valid street address differed by treatment condition.) Children without valid street addresses were excluded from the analyses. Importantly, as described below, this study adjusts for differential selection into the respective analytic samples based on the probability of having a valid home address. And as described in more detail in the Results section, results were robust to the exclusion of these sampling weights, indicating that bias due to having a valid street address was unlikely.

Measures

Baseline Characteristics

Data on baseline descriptive characteristics for the full RCT sample were gathered from the state administrative database. These baseline descriptive characteristics include *date of birth, a binary indicator for female, race/ethnicity, age*, and *English as a primary language*. In addition to baseline data gathered from the state, children in the ISS sample were administered a series of assessments and questionnaires during the fall of their prekindergarten year that was used to create a more robust set of baseline controls for the ISS sample. The baseline parent questionnaire provided indicators of *mother's education, number of working parents, number of household magazine and newspaper subscriptions*, and *the frequency with which families used the library*. Baseline achievement was also assessed with a selection of scales from the Woodcock Johnson III Achievement Battery (Woodcock, McGre, & Mather, 2001). These scales included one measure of reading (Passage Comprehension), three measures of math (Applied Problem Solving, Quantitative Concepts, and Calculation), two measures of language (Oral Comprehension and Picture Vocabulary), and two measures of literacy (Letter-Word Identification and Spelling). For all baseline covariates (excluding neighborhood poverty), multiple imputation was used to fill in missing values due to item-specific nonresponse. Missing data on baseline covariates were rare for both the RCT and ISS samples, with rates ranging from 0.0% to 1.2% in the RCT sample and from 0.0% to 5.5% in the ISS sample. (A complete list of missing data patterns is provided in the Supplementary Appendix Table G.1.) Subsequent analyses were based on 25 multiply imputed datasets computed separately for each analytic sample that were combined based on Rubin's (1987) rules.

Neighborhood Poverty

Although neighborhood disadvantage can be measured using a variety of indicators, this study focuses on neighborhood poverty because prior research suggests that neighborhood poverty is closely associated with underlying social processes believed to be responsible for neighborhood effects (Wilson, 2012). Moreover, neighborhood poverty, unlike multidimensional scales of neighborhood disadvantage, has a straightforward, policy-relevant interpretation that is not reliant on distributional considerations (Wodtke, 2013). In particular, prior neighborhood effect research has noted the existence of thresholds, or "tipping points," after which neighborhoods begin to matter in shaping children's developmental outcomes (Galster, 2012).

This study uses as its primary measure of neighborhood poverty a three-level ordinal measure (coded 1, 2, and 3), to indicate whether a child lived at baseline in a low-poverty neighborhood (below 10% poor), moderate-poverty neighborhood (10% to 30% poor), or high-poverty neighborhood (at least 30% poor). Although thresholds of neighborhood poverty used in prior literature are somewhat varied (see D. J. Harding, 2003; Jargowsky, 1997; Wodtke, 2013), exploratory analysis indicated that measures based on the 10% and 30% thresholds best captured the relation between neighborhood poverty and preschool effects.

Measures of neighborhood poverty were based on block-group-level data gathered from the 2008–2012 American Community Survey. Block groups are standard geographical units made available by the U.S. Census Bureau that are composed of contiguous clusters of residential blocks that contain between 600 and 3,000 people. The boundaries of block groups are mostly defined by local participants in the Census Bureau's Participant Statistical Areas Program, suggesting that boundaries are oftentimes locally meaningful and reflective of residential divisions (Bureau of the Census, 2008). (As in the Results section, Supplementary Appendix Table B examines the robustness of estimates to alternative thresholds of neighborhood poverty, various measures of neighborhood disadvantage including a multidimensional composite index, as well as different conceptions of neighborhood boundaries.)

Outcome

The Tennessee Comprehensive Assessment Program (TCAP) requires students attending public school in Tennessee to be tested annually in core subject areas from third grade through eighth grades. The outcome variables in the present study were children's *reading* and *math* performance, respectively, on these statewide achievement tests in third grade. Like baseline data for the full sample, these achievement data were gathered from the state educational database. It is important to note that some children in the second cohort were retained and had not yet reached third grade by the time follow-up data collection occurred. However, as indicated in the Supplementary Appendix Table D.1, there were no statistically significant differences in retention across treatment conditions in any level of neighborhood poverty in either the RCT or ISS sample. Therefore, retention differences should not bias treatment contrasts among children who did take the test in third grade.

Weighting

As noted previously, not every child in the RCT or ISS samples provided a valid street address at baseline that could be geolocated. In addition, not every child took the TCAP in third grade. To correct for potential nonrandom "attrition," this study used a poststratification weighting technique to adjust for differential probabilities of being included in the analytic samples. In effect, children more likely to misreport a valid street address or to have a missing TCAP score in third grade were up-weighted relative to their peers. These sampling weights were estimated from two logistic regressions in which a binary indicator of whether a child had a valid street address at baseline or whether a child took the TCAP in third grade, respectively, was regressed on the full set of sample-specific baseline characteristics as well as a set of subsequent child-level variables gathered from the state administrative database. Results from these regressions (and all included covariates) are provided in the Supplementary Appendix Table F.1. The inverse of these two predicted probabilities became the poststratification address and TCAP weights, respectively.

Effect estimates reported below for the full RCT sample were weighted by the product of the address weight and the TCAP weight. Effect estimates for the ISS sample were weighted similarly, except for the addition of the consent weight described earlier. That is, effect estimates for the ISS sample were weighted by the product of all three weights (see Wodtke, Harding, & Elwert, 2011 for similar weighting technique). As described in the Results section, results were robust to the exclusion of each of these weights, indicating that bias due to missing a valid street address or missing outcome data—or, for the ISS sample, bias due to differential consent rates—was unlikely.

Balance Tests and Summary Statistics

Although randomization into treatment conditions eliminates observed and unobserved differences across treatment and control groups in expectation, randomization does not necessarily ensure randomization within subgroups defined by baseline characteristics (VanderWeele & Knol, 2011). Moreover, it is unclear whether the extent to which any nonrandom "attrition" based on misreporting of baseline addresses or incomplete outcome data may have affected randomization. Therefore, this section reports a series of balance tests across the RCT and ISS samples, respectively, to determine the extent of subgroup balance on baseline covariates. Unadjusted differences were assessed using ordinary least squares regressions with ITT condition as the only predictor. Adjusted differences were based on regressions that included the other baseline characteristics germane to each analytic sample as covariates. Adjusted models included indicators for preK center. All models used sampling weights.

Tables 1 and 2 present unadjusted and adjusted comparisons of baseline characteristics between treatment and control groups within each level of neighborhood poverty for the RCT and ISS samples, respectively. As shown in Columns (3), (6), and (9) of each table, the few covariate imbalances that were present in unadjusted comparisons were considerably reduced in the fully adjusted comparisons. In the RCT sample (Table 1), 1 of the 24 comparisons was significant with $p < .10$, and 1 of the 24 comparisons was significant with $p < .05$ based on t tests that did not adjust for multiple comparisons; in the ISS sample (Table 2), 3 of the 57 comparisons were significant with $p < .10$, and 2 of the 57 comparisons were significant with $p < .05$ based on analogous t tests that did not adjust for multiple comparisons. These patterns are generally consistent with what one would expect under random assignment (i.e., simply by chance, one would expect 1 out of 20 and 1 out of 10 comparisons to show up as significant at the $p < .05$ and $p < .10$ levels, respectively).

Table 1

Summary Statistics and Balance Tests for Children in TN-VPK-Neighborhood Data Linked Sample Across Levels of Neighborhood Poverty, Full Sample

	Levels of Neighborhood Poverty								
	Low Poverty (<10% Poor)			Moderate Poverty (10% to 30% Poor)			High Poverty (>30% Poor)		
	Control Group, *M* (1)	Experimental Versus Control (2)	Adjusted Difference (3)	Control Group, *M* (4)	Experimental Versus Control (5)	Adjusted Difference (6)	Control Group, *M* (7)	Experimental Versus Control (8)	Adjusted Difference (9)
Available test score	0.84	0.01 (0.03)	−0.01 (0.03)	0.85	0.03 (0.02)	0.02 (0.03)	0.86	−0.02 (0.04)	−0.02 (0.04)
Female	0.51	−0.02 (0.04)	0.02 (0.04)	0.49	0.02 (0.03)	0.06 (0.04)	0.48	0.07 (0.06)	0.04 (0.07)
Black	0.45	−0.08 (0.06)	0.01 (0.03)	0.61	−0.15* (0.07)	−0.02 (0.03)	0.75	0.02 (0.11)	0.01 (0.03)
Hispanic	0.20	−0.06* (0.03)	−0.01 (0.01)	0.27	−0.07* (0.03)	0.00 (0.01)	0.33	−0.11* (0.05)	0.01 (0.02)
White	0.54	0.10* (0.04)	−0.01 (0.02)	0.39	0.14*** (0.04)	−0.02 (0.01)	0.28	0.09 (0.05)	−0.00 (0.02)
English language learner	0.76	0.07* (0.03)	−0.04* (0.02)	0.68	0.13*** (0.03)	0.03 (0.02)	0.64	0.14** (0.05)	0.04† (0.02)
Age	4.90	0.22*** (0.04)	0.03 (0.02)	4.97	0.09* (0.04)	−0.02 (0.03)	4.97	0.17** (0.06)	−0.03 (0.03)
Neighborhood poverty	0.05	0.00 (0.00)	−0.00 (0.00)	0.18	0.01 (0.01)	−0.00 (0.00)	0.42	−0.01 (0.02)	−0.01 (0.01)
N	395	513		452	513		169	198	

Note. TN-VPK = Tennessee Voluntary Prekindergarten Program. This table presents summary statistics and balance tests of equivalency for a subset of variables collected prior to randomization. The first row indicates the share of children living in each level of neighborhood poverty that took the state-wide achievement assessment in third grade. The remaining rows refer to the sample of children who had valid achievement data in third grade and who provided valid street addresses at baseline. Columns (1), (2), and (3) include children living in low-poverty neighborhoods; Columns (4), (5), and (6) include children living in moderate-poverty neighborhoods; Columns (7), (8), and (9) refer to children living in high-poverty neighborhoods. Columns (1), (4), and (7) show the control group mean for each variable across quartiles of neighborhood disadvantage. Columns (2), (5), and (8) report unadjusted differences between the experimental VPK group and the control group, which is estimated using a weighted ordinary least squares regression of each variable on a binary indicator for random assignment into the treatment condition. The adjusted comparisons in Columns (3), (6), (9), and (12) were based on regressions that included as covariates the complete set of baseline characteristics. Adjusted comparisons also included randomization pool fixed effects. All comparisons were weighted to adjust for children's differential probability of (1) having taken the state-wide achievement test in third grade and (2) having reported a valid street address at baseline. Standard errors are reported in parenthesis and are clustered at the neighborhood level.

†$p < .10$. *$p < .05$. **$p < .01$. ***$p < .001$, for two-tailed tests of significance.

Table 2

Summary Statistics and Balance Tests for Children in TN-VPK-Neighborhood Data Linked Sample Across Levels of Neighborhood Poverty, ISS Sample

	Levels of Neighborhood Poverty								
	Low Poverty (<10% Poor)			Moderate Poverty (10% to 30% Poor)			High Poverty (>30% Poor)		
	Control Group, *M* (1)	Experimental Versus Control (2)	Adjusted Difference (3)	Control Group, *M* (4)	Experimental Versus Control (5)	Adjusted Difference (6)	Control Group, *M* (7)	Experimental Versus Control (8)	Adjusted Difference (9)
Available test score	0.78	−0.01 (0.05)	0.00 (0.00)	0.84	0.05 (0.04)	0.00 (0.00)	0.74	0.13 (0.09)	0.00 (0.00)
Female	0.49	0.03 (0.08)	0.12 (0.12)	0.43	0.06 (0.07)	0.08 (0.08)	0.54	0.07 (0.15)	−0.19 (0.30)
Black	0.39	−0.02 (0.13)	−0.07 (0.06)	0.61	−0.11 (0.15)	−0.00 (0.11)	0.80	−0.01 (0.25)	−0.11 (0.15)
Hispanic	0.23	−0.09 (0.07)	−0.04 (0.03)	0.27	$-0.10^{\dagger}$ (0.06)	0.06 (0.04)	0.36	−0.02 (0.11)	0.20* (0.10)
White	0.55	0.02 (0.08)	$-0.16^{\dagger}$ (0.09)	0.39	$0.14^{\dagger}$ (0.07)	−0.06 (0.07)	0.14	0.06 (0.08)	−0.04 (0.06)
English language learner	0.72	0.11 (0.07)	−0.07 (0.06)	0.60	0.25*** (0.07)	0.13** (0.04)	0.58	0.08 (0.12)	0.08 (0.12)
Age	4.96	$0.12^{\dagger}$ (0.07)	−0.05 (0.06)	5.02	0.03 (0.08)	0.04 (0.04)	4.90	0.13 (0.16)	−0.03 (0.08)
Letter-word	318.79	0.61 (4.28)	1.76 (4.15)	315.11	5.89 (3.84)	3.30 (3.17)	317.86	−4.91 (7.01)	6.71 (6.07)
Spelling	351.99	−1.51 (3.94)	−2.17 (4.95)	354.09	−1.85 (4.03)	−0.24 (4.17)	354.03	−8.76 (5.51)	−0.60 (13.80)
Oral comprehension	440.96	3.29 (2.36)	0.10 (1.82)	440.71	3.37 (2.63)	−2.71 (2.36)	438.47	3.06 (4.18)	0.37 (4.79)
Picture vocabulary	451.77	5.81 (3.78)	0.49 (2.76)	448.64	8.36* (4.04)	−1.32 (2.07)	443.20	5.41 (8.88)	2.27 (5.61)
Applied problem solving	387.14	6.63 (4.03)	4.11 (3.29)	386.20	7.44* (3.42)	1.42 (3.25)	385.41	−1.32 (7.73)	7.16 (4.78)
Quantitative concept	406.48	0.39 (1.80)	−0.04 (2.17)	408.00	−0.11 (2.06)	−2.35 (1.79)	408.45	−4.56 (3.96)	$-8.47^{\dagger}$ (4.85)
Mother high school noncompleter	0.09	0.05 (0.04)	−0.14 (0.12)	0.19	−0.03 (0.05)	0.18 (0.13)	0.22	0.00 (0.13)	−0.14 (0.32)
% with two working parents	0.35	0.02 (0.07)	−0.10 (0.15)	0.35	−0.03 (0.06)	−0.04 (0.12)	0.24	0.00 (0.10)	$0.33^{\dagger}$ (0.19)
Library use	0.45	0.01 (0.07)	0.11 (0.18)	0.42	−0.08 (0.07)	−0.01 (0.12)	0.60	−0.09 (0.15)	0.35 (0.23)

(continued)

Table 2 **(continued)**

	Levels of Neighborhood Poverty								
	Low Poverty (<10% Poor)			Moderate Poverty (10% to 30% Poor)			High Poverty (>30% Poor)		
	Control Group, *M* (1)	Experimental Versus Control (2)	Adjusted Difference (3)	Control Group, *M* (4)	Experimental Versus Control (5)	Adjusted Difference (6)	Control Group, *M* (7)	Experimental Versus Control (8)	Adjusted Difference (9)
Newspaper subscriptions	0.86	−0.02 (0.05)	−0.03 (0.11)	0.80	0.03 (0.06)	−0.03 (0.15)	0.85	−0.00 (0.10)	−0.14 (0.30)
Magazine subscriptions	0.76	−0.02 (0.05)	0.00 (0.09)	0.75	−0.02 (0.06)	−0.16 (0.10)	0.85	−0.12 (0.11)	−0.14 (0.22)
Neighborhood poverty	0.04	0.00 (0.01)	0.00 (0.00)	0.18	0.01 (0.01)	0.01 (0.01)	0.43	−0.03 (0.03)	−0.02 (0.03)
N	110	169		156	205		37	55	

Note. TN-VPK = Tennessee Voluntary Prekindergarten Program. ISS = Intensive Substudy. This table presents summary statistics and balance tests of equivalency for a subset of variables collected prior to randomization. The first row indicates the share of children living in each level of neighborhood poverty that took the state-wide achievement assessment in third grade. The remaining rows refer to the sample of children who had valid achievement data in third grade and who provided valid street addresses at baseline. Columns (1), (2), and (3) include children living in low-poverty neighborhoods; Columns (4), (5), and (6) include children living in moderate-poverty neighborhoods; Columns (7), (8), and (9) refer to children living in high-poverty neighborhoods. Columns (1), (4), and (7) show the control group mean for each variable across quartiles of neighborhood disadvantage. Columns (2), (5), and (8) report unadjusted differences between the experimental VPK group and the control group, which is estimated using a weighted ordinary least squares regression of each variable on a binary indicator for random assignment into the treatment condition. The adjusted comparisons in Columns (3), (6), (9), and (12) were based on regressions that included as covariates the complete set of baseline characteristics. Adjusted comparisons also included randomization pool fixed effects. All comparisons were weighted to adjust for children's differential probability of (1) having taken the state-wide achievement test in third grade and (2) having reported a valid street address at baseline. Standard errors are reported in parenthesis and are clustered at the neighborhood level.

$^{\dagger}p < .10$. $^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$, for two-tailed tests of significance.

Although covariate adjustment yielded statistically similar treatment conditions across nearly all the baseline characteristics in the three subgroups of neighborhood poverty, this approach may still be inadequate to account for bias due to unobserved confounding, which remains the primary threat to the internal validity of the present study (Rosenbaum & Rubin, 1983). In subsequent analytic models, each of the specified baseline characteristics was included as a covariate to improve precision and to adjust for any initial biases that may have been introduced by the compromised randomization inherent to a post hoc subgroup analysis with obvious attrition due to misreporting of baseline addresses and incomplete outcome data.

The bottom rows in Tables 1 and 2 report the distribution of the analytic samples across levels of neighborhood poverty. With regard to the RCT sample, 16.3% of children lived at baseline in high-poverty neighborhoods. The largest share of children in the RCT sample lived in moderate-poverty neighborhoods (43.1%), while similar percentages lived in low-poverty neighborhoods (40.5%). These patterns are similar with respect to the ISS sample, with the distinction being that a slightly larger share of children (49.1%) lived in the moderate-poverty neighborhoods, 38.6% in low-poverty neighborhoods, and 12.3% in high-poverty neighborhoods.

Analysis

This study estimates moderated ITT effects of the VPK treatment across levels of neighborhood poverty. These estimates are essentially covariate-adjusted comparisons of control and treatment group means across levels of neighborhood poverty. Following prior research on variation in treatment effects (e.g., Schochet & Deke, 2014), this study estimates moderated ITT effects of preK on children's third-grade reading and math achievement (y), respectively, using ordinary least squares regression specification of the following form:

$$y_i = \alpha + VPK_i\left(\theta + \sum_{l=1}^{3} Z_i NP_{li}\right) + \sum_{j=1}^{K} X_{ji}\gamma_j + s_i\delta + \varepsilon_i \quad (1)$$

where *VPK* is an indicator variable for being randomly assigned to VPK, *NP* is a categorical measure of neighborhood poverty coded 1 through 3 that records whether children resided at baseline in a low-poverty neighborhood (below 10% poor), moderate-poverty neighborhood (at least 10% poor but less than 30% poor), or high-poverty neighborhood (at least 30% poor). Mechanically, *NP* is arrayed as series of binary indicators for each level of neighborhood poverty with low-poverty neighborhoods serving as the referent category.[2] *X* represents a vector of sample-specific baseline covariates included to improve precision and account for chance differences between groups in the distribution of baseline characteristics. Finally, *s* is a set of indicators for randomization site that ensure that third-grade achievement patterns were not associated with any unobserved characteristics that systematically varied across preK classrooms. All regressions were weighted to adjust for differential probabilities of being included in the sample as described previously. Standard errors were clustered at the neighborhood level.

Two assumptions are required for moderated effects to be interpreted causally. Although neighborhoods were not randomized in the study, a causal interpretation can be made if it is assumed that (a) randomization yielded subgroup randomization within levels of neighborhood poverty and (b) compliance rates did not vary across these subgroups. As discussed previously, Tables 1 and 2 provided no evidence of covariate imbalance across levels of neighborhood poverty. Moreover, Supplementary Appendix Table E.1 provides no evidence that compliance rates varied across levels of neighborhood poverty. In short, there is no evidence to discourage a causal interpretation of moderated effects. Thus, the estimate of θ in Equation (1) can be interpreted as identifying the causal impact of being offered admission into VPK for children in low-poverty neighborhoods, while estimates of Z_2 and Z_3 in Equation (1) identify the increment (or decrement) to the causal impact of being offered VPK admission for those in moderate- and high-poverty neighborhoods, respectively. It is also important to note that the inclusion of fixed effects for randomization site (s) controls for unobserved differences across preK classrooms and means that results are interpreted as within-randomization-list estimates, that is, as the potential differential effect of preK across neighborhood contexts for children within the same randomization list. To facilitate interpretation of results, this study also reports results in standard deviation differences, which are readily interpreted as effect sizes. Effect sizes were computed by dividing the coefficient of interest by a weighted average of the pooled standard deviation across levels of neighborhood poverty.

Because of the potential of crossovers, that is, some children offered admission into VPK did not enroll in VPK and some children assigned to the control group wound up attending VPK anyway, these ITT estimates are likely different from the causal effect of actually enrolling in VPK. (Supplementary Appendix Table E.1 shows compliance rates across categories of neighborhood poverty.) This study follows Chetty et al. (2015) and estimates the moderated impact of VPK enrollment—the moderated local average treatment effect (LATE)—by leveraging a two-stage least squares approach and instrumenting for VPK enrollment and the interaction between VPK enrollment and the three-level neighborhood poverty variable with an indicator for treatment assignment and interactions between treatment assignment and the three-level neighborhood poverty variable. Formally, this specification takes the following form:

$$y_i = \alpha + EnrollVPK_i\left(\theta + \sum\nolimits_{l=1}^{3} Z_i NP_{li}\right) + \sum\nolimits_{j=1}^{K} X_{ji}\gamma_j + s_i\delta + \varepsilon_i \quad (2)$$

where *EnrollVPK* and the interactions between *EnrollVPK* and *NP* are the instrumented indicators for enrollment in VPK and its interaction with neighborhood poverty. The remainder of components, clustering, and weighting strategy are identical to those in Equation (1). Under the assumptions that VPK assignment only affects outcomes through VPK enrollment and that subgroup randomization holds within levels of neighborhood poverty, θ can be interpreted as the causal effect of VPK enrollment for children living in low-poverty neighborhoods, while Z_2 and Z_3 can be interpreted as the change in the causal effect of VPK enrollment for children living in moderate- and high-poverty neighborhoods relative to those living in low-poverty neighborhoods.

In addition to examining moderated ITT and LATE effects of VPK across levels of neighborhood poverty, this study is also interested in potential explanations for why VPK effects might differ across levels of neighborhood poverty. This study sheds light on this issue by examining the extent to which factors correlated with neighborhood poverty—*rival moderators*—might account for why neighborhood poverty matters for differential VPK effects. This study focuses on four such rival moderators described in the background section of this article: differential access to alternative childcare options, contextual risk factors, differential sorting of families by SES, and urbanicity.[3] In particular, this *rival moderator analysis* is based on models that take the following form:

$$y_i = \alpha + VPK_i\left(\theta + \sum\nolimits_{l=1}^{3} Z_i NP_{li}\right) + \delta(VPK_i \times AltMod_i) + \sum\nolimits_{j=1}^{K} X_{ji}\gamma_j + s_i\delta + \varepsilon_i \quad (3)$$

$$y_i = \alpha + EnrollVPK_i\left(\theta + \sum\nolimits_{l=1}^{3} Z_i NP_{li}\right) + \delta(EnrollVPK_i \times AltMod_i) + \sum\nolimits_{j=1}^{K} X_{ji}\gamma_j + s_i\delta + \varepsilon_i. \quad (4)$$

Equations (3) and (4) are identical to Equations (1) and (2) except that Equations (3) and (4) add an interaction between VPK and the rival moderator(s) of interest ($AltMod_i$). [The components, weighting, and clustering in Equations (3) and (4) are identical to those described for Equations (1) and (2).] Notably, for the LATE estimates described in Equation (4), *EnrollVPK* and the interactions between *EnrollVPK* and the moderators of interest were instrumented with an indicator for treatment assignment and interactions between treatment assignment and moderators of interest. Of interest in Equations (3) and (4) is the change in the coefficient for the interaction between preK and neighborhood poverty after the inclusion of each moderator(s) of interest compared with the coefficient for the interaction between preK and neighborhood poverty relative to Equations (1) and (2). It is important to underscore that this rival moderator analysis is purely exploratory and should not be interpreted in causal terms because there are likely a number of unobserved neighborhood- and family-level characteristics that correlate with the rival moderators of interest.

The following section reports estimates of Equations (1) and (2) for children's third-grade reading and math achievement, respectively, before turning to the rival moderator analysis (Equations 3 and 4). For reference, the following section also reports main effect estimates of ITT and LATE that exclude interaction terms from Equations (1) and (2). The coefficients of interest in these main effect equations are interpreted as the average effect of VPK assignment and enrollment, respectively, across the entire sample of participants, controlling for neighborhood poverty. Also, given that this study is probing interactions that the original TN-VPK experiment was not set up to answer, the analyses have limited power to detect differential effects. Thus, a more flexible significance level is used throughout the Results section ($p < .10$) than is typical in preK literature.

Results

This section begins by describing results for the full sample and concludes with results for the ISS sample. Overall, this study finds considerable evidence that the estimated effect of VPK on reading achievement depended on the residential environments in which children lived at baseline, with significant positive effects being observed for those children living in high-poverty neighborhoods but negative effects for children living in low-poverty neighborhoods. Table 3 reports full sample estimates of the effect of VPK on children's reading and math achievement in third grade, respectively. Table 4 reports analogous estimates for the ISS sample. Columns (1) through (4) refer to reading achievement; Columns (5) through (8) refer to math achievement. As noted above, the primary focus of the discussion concerns effect sizes because of their ease of interpretation: Effect sizes are readily interpreted as a standard deviation difference. Effect sizes are provided in brackets below the coefficient and standard errors in Tables 3 and 4.

Full Sample: Reading Achievement

Column (1) of Table 3, which reports ITT estimates for the full sample, provides no evidence that, on average, random assignment into VPK classrooms had an effect on third-grade reading achievement. However, Column (2) in Table 3 shows that the estimated effect of VPK on third-grade reading achievement varied across levels of neighborhood poverty. In particular, the point estimate for the main effect of VPK in Column (2) indicates that, among children living in low-poverty neighborhoods, children assigned to VPK scored 0.13 standard deviations (*SD*s) lower than those not assigned to VPK in reading achievement in third grade ($\beta = -4.44$, $p = .067$). In

Table 3

Moderated Effect of Tennessee Voluntary PreK on Children's Third-Grade Achievement Across Levels of Neighborhood Poverty at Baseline, Full Sample Results

	Reading Achievement (n = 2,240)				Math Achievement (n = 2,239)			
	ITT Estimates		TOT Estimates		ITT Estimates		TOT Estimates	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
VPK	−1.51	−4.44†	−3.84	−10.02†	−3.32†	−3.91	−8.43†	−9.59†
	(1.82)	(2.42)	(4.51)	(5.33)	(1.75)	(2.65)	(4.40)	(5.80)
	[−0.04]	[−0.13]	[−0.11]	[−0.29]	[−0.10]	[−0.11]	[−0.25]	[−0.28]
Moderate poverty		−2.31		−3.07		−2.15		−0.16
		(2.58)		(4.91)		(2.57)		(4.99)
		[−0.07]		[−0.09]		[−0.06]		[−0.00]
High poverty		−7.72*		−18.96**		−4.76		−10.11†
		(3.24)		(6.68)		(3.18)		(5.73)
		[−0.22]		[−0.55]		[−0.14]		[−0.29]
VPK × Moderate Poverty		1.61		3.02		−1.33		−3.37
		(3.34)		(6.82)		(3.48)		(7.15)
		[0.05]		[0.09]		[−0.04]		[−0.10]
VPK × High Poverty		13.53**		28.68**		6.94†		14.59†
		(4.70)		(9.90)		(4.21)		(8.65)
		[0.39]		[0.83]		[0.20]		[0.42]
R^2	0.11	0.12	0.11	0.10	0.11	0.11	0.10	0.09

Note. ITT = intent-to-treat; TOT = treatment-on-the-treated; VPK = Voluntary Prekindergarten Program. Columns (1) through (4) refer to children's reading achievement in third grade. Columns (5) through (8) refer to children's math achievement in third grade. Columns (1) and (5) report ITT estimates from an ordinary least squares regression of children's achievement on an indicator for random assignment into the experimental preK group. Columns (2) and (6) interact the indicator of random assignment with a measure of neighborhood poverty at baseline. Columns (3) and (7) report LATE estimates using a two-stage least squares specification, instrumenting for preK enrollment with the experimental indicator of preK assignment. Columns (4) and (8) instrument for preK enrollment and the interaction of preK enrollment and levels of neighborhood poverty with the experimental indicator of preK assignment and the interaction of preK assignment and levels of neighborhood poverty. Moderate-poverty neighborhoods had poverty rates that met or exceeded 10% but were below 30%. High poverty neighborhoods had poverty rates that met or exceeded 30%. The referent category in the table is low-poverty neighborhoods, which had poverty rates below 10%. All models included randomization pool fixed effects and were weighted to adjust for differences in children's likelihood of having a nonmissing third-grade achievement score and a valid street address at baseline. All models controlled for baseline characteristics at the individual and neighborhood level. Individual-level controls included age, race, gender, primary language, and cohort. Neighborhood poverty rates were measured at the block-group level and were gathered from the 2008-2012 American Community Survey. Standard errors are in parenthesis and are clustered at the neighborhood level. Effect sizes are in brackets. Effect sizes were computed by dividing the coefficient of interest by a weighted average of the pooled standard deviation across levels of neighborhood poverty.
$^{\dagger}p < .10$. $^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$, for two-tailed tests of significance.

Table 4

Moderated Effect of Tennessee Voluntary PreK on Children's Third Grade Achievement Across Neighborhood Poverty Levels at Baseline, Intensive Substudy Sample Results

	Reading Achievement (n = 732)				Math Achievement (n = 732)			
	ITT Estimates		TOT Estimates		ITT Estimates		TOT Estimates	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
VPK	−1.67	−4.00	−3.30	−6.13	−5.82	−6.76	−11.50†	−10.90
	(3.35)	(4.82)	(6.14)	(6.89)	(3.58)	(6.83)	(6.67)	(9.71)
	[−0.05]	[−0.12]	[−0.10]	[−0.18]	[−0.16]	[−0.19]	[−0.32]	[−0.30]
Moderate poverty		0.82		4.89		−1.97		3.67
		(5.28)		(8.82)		(7.68)		(11.84)
		[0.02]		[0.15]		[−0.05]		[0.10]
High poverty		−12.77†		−38.67†		−13.47		−25.29
		(7.63)		(21.13)		(9.02)		(15.88)
		[−0.38]		[−1.16]		[−0.37]		[−0.70]
VPK × Moderate Poverty		−2.05		−6.24		−2.48		−8.54
		(6.03)		(10.15)		(8.51)		(13.40)
		[−0.06]		[−0.19]		[−0.07]		[−0.24]
VPK × High Poverty		22.31*		51.45*		14.19		27.26
		(9.25)		(25.03)		(9.50)		(18.08)
		[0.67]		[1.55]		[0.39]		[0.75]
R^2	0.32	0.33	0.32	0.28	0.30	0.30	0.29	0.29

Note. ITT = intent-to-treat; TOT = treatment-on-the-treated; VPK = Voluntary Prekindergarten Program. Columns (1) through (4) refer to children's reading achievement in third grade. Columns (5) through (8) refer to children's math achievement in third grade. Columns (1) and (5) report ITT estimates from an ordinary least squares regression of children's achievement on an indicator for random assignment into the experimental preK group. Columns (2) and (6) interact the indicator of random assignment with a measure of neighborhood disadvantage at baseline. Columns (3) and (7) report local average treatment effect estimates using a two-stage least squares specification, instrumenting for preK enrollment with the experimental indicator of preK assignment. Columns (4) and (8) instrument for preK enrollment and the interaction of preK enrollment and neighborhood disadvantage with the experimental indicator of preK assignment and the interaction of preK assignment and neighborhood disadvantage. All models included randomization pool fixed effects and were weighted to adjust for differences in children's likelihood of having a nonmissing third-grade achievement score and a valid street address at baseline. All models controlled for baseline characteristics at the individual and neighborhood level. Individual-level controls included age, race, gender, primary language, cohort, mother's education, number of working parents, number of household magazine and newspaper subscriptions, frequency of library usage, and children's baseline achievement on six cognitive assessments: Letter Word, Applied Problem Solving, Quantitative Concepts, Story Recall, Picture Vocabulary, and Oral Comprehension. Neighborhood poverty rates were measured at the block-group level and were gathered from the 2008-2012 American Community Survey. Moderate-poverty neighborhoods had poverty rates that met or exceeded 10% but were below 30%. High-poverty neighborhoods had poverty rates that met or exceeded 30%. The referent category in the table is low-poverty neighborhoods, which had poverty rates below 10%. Standard errors are in parenthesis and are clustered at the neighborhood level. Effect sizes are in brackets. Effect sizes were computed by dividing the coefficient of interest by a weighted average of the pooled standard deviation across levels of neighborhood poverty.

$^{\dagger}p < .10$. $^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$, for two-tailed tests of significance.

contrast, the point estimate for the interaction in Column 2 between VPK assignment and high-poverty neighborhood indicates that the effect of VPK assignment was 0.39 *SD*s larger for children living at baseline in high- compared with low-poverty neighborhoods (β = 13.53, p = .004).

This general pattern is echoed in Columns (3) and (4) of Table 3, which report LATE estimates of the main and moderated effect of VPK on third-grade reading achievement, respectively. Similar to the ITT estimates, Column (3) provides no evidence that VPK *enrollment* had an effect, on average, on third-grade reading achievement. However, Column (4) shows heterogeneity across levels of neighborhood poverty. The significant point estimate for VPK in Column (4) shows that children living at baseline in low-poverty neighborhoods who enrolled in VPK scored −0.29 *SD*s lower in reading achievement than their respective control group in third grade (β = −10.02, p = .060). However, the significant interaction term in Column 4 between VPK enrollment and high-poverty neighborhood indicates that the estimated difference in achievement between treatment and control group children was 0.83 *SD*s larger (more positive) for children living at baseline in high- compared with low-poverty neighborhoods (β = 28.68, p = .004).

Although the significant interaction terms provide evidence that the effect of VPK differs for children living in high- compared with low-poverty neighborhoods, what remains unclear is whether children living in high-poverty neighborhoods actually experienced a positive effect from VPK. Figures 1 and 2 plot the average marginal effect of VPK on third-grade reading achievement (y-axis) across levels of neighborhood poverty (x-axis). These average marginal effects are equivalent to the adjusted difference in means between treatment conditions within each level of neighborhood poverty. The panel on the left of each figure refers to the RCT sample; the panel on the right refers to the ISS sample, discussed below. The points in each figure specify the average marginal effect of VPK within each level of neighborhood poverty. The bars correspond to 90% confidence intervals.

Overall, Figures 1 and 2 show that the effect of VPK on third-grade reading achievement for the full RCT sample was positive and statistically meaningful for children in high-poverty neighborhoods at baseline. Among this subset, the average marginal effect of VPK *assignment* on third-grade reading achievement was 0.27 *SD*s (p = .029); the average marginal effect of VPK *enrollment* on third-grade reading achievement was 0.54 *SD*s (p = .041). In contrast, point estimates in the far left of the panel in each figure, illustrating average marginal effects of VPK for children in low-poverty neighborhoods, reiterate the earlier noted finding that both assignment to and enrollment in VPK exerted a negative effect on third-grade reading achievement in the full sample (p = .067 and p = .060, respectively). No evidence supports either VPK assignment or enrollment having an effect on third-grade achievement for children living in moderate-poverty neighborhoods at baseline.

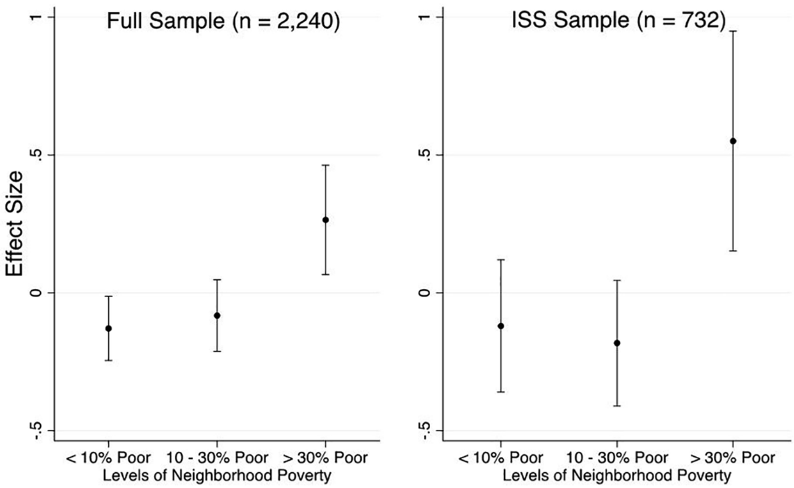

Figure 1. **ITT estimates of the effect of VPK on third-grade reading achievement across levels of neighborhood poverty.**

Note. ITT = intent to treat; VPK = Voluntary Prekindergarten program; ISS = Intensive Substudy.

Full Sample: Math Achievement

Columns (5) to (8) in Table 3 report results for third-grade math achievement in the RCT sample. In contrast to the null main effects observed for reading achievement, this study finds some evidence that the control group, on average, outperformed the treatment group on math achievement in third grade. Column 5 in Table 3 provides evidence that, on average, children in the *control group* (i.e., those randomly assigned not to attend VPK) outperformed children randomly assigned to attend VPK by 0.10 *SD*s ($\beta = -3.32$, $p = .055$). Similarly, those *enrolled* in VPK scored 0.25 *SD*s lower on third-grade math achievement, on average, than the control group ($\beta = -8.43$, $p = .055$), as shown in Column (7). Similar to that reported for reading achievement, Columns (6) and (8) show that the effects of VPK varied across levels of neighborhood poverty. In particular, Column (6) shows that the effect of VPK assignment on children's third-grade math achievement was 0.20 *SD*s larger for children in high- compared with low-poverty neighborhoods ($\beta = 6.94$, $p = .098$). Similarly, Column (8) shows that the effect of VPK enrollment on children's third-grade math achievement was 0.42 *SD*s larger for children in high- compared with low-poverty neighborhoods ($\beta = 14.59$, $p = .090$).

Despite observing that VPK effects differed across levels of neighborhood poverty, the left panels of Figures 3 and 4 provide no evidence that

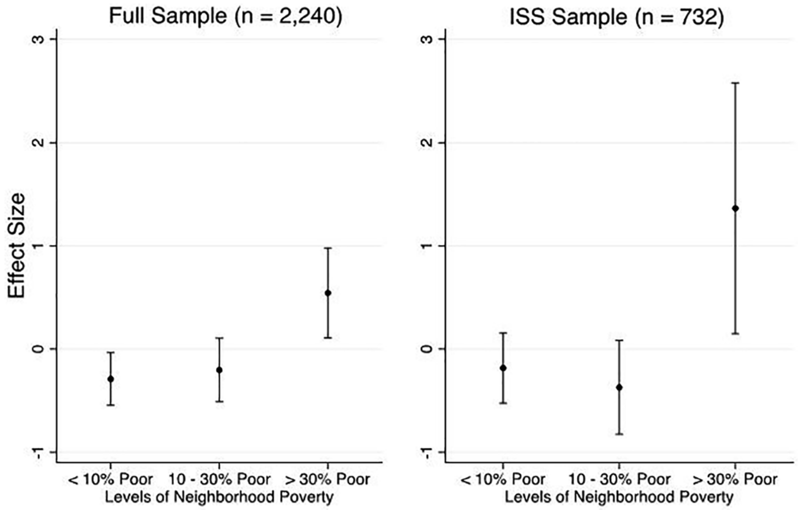

Figure 2. **LATE estimates of the effect of VPK on third-grade reading achievement across levels of neighborhood poverty.**

Note. LATE = local average treatment effect; VPK = Voluntary Prekindergarten program; ISS = Intensive Substudy.

VPK affected third-grade math achievement for children living in high-poverty neighborhoods. Said otherwise, there is no evidence that the effect of VPK for children living in high-poverty neighborhoods was statistically different from zero despite being significantly larger than the observed effect for children living in low-poverty neighborhoods. However, Figures 3 and 4 do provide evidence that among children living in moderate-poverty neighborhoods, those assigned to the control group outperformed their peers assigned to the treatment group on third-grade math achievement (p = .048), while children who did not enroll in VPK outperformed their peers who did enroll in VPK on third-grade math achievement in low- (p = .098) and moderate-poverty neighborhoods, respectively (p = .042).

Intensive Substudy Sample

Results for the ISS sample in Table 4 are arranged similarly to Table 3: Columns (1) to (4) report results for reading achievement; Columns (5) to (8) report results for math achievement; effect sizes are in brackets. Substantive conclusions for the ISS sample are generally consistent with those reported in the previous section. Based on a subset of the full sample for whom baseline achievement and socioeconomic variables were available,

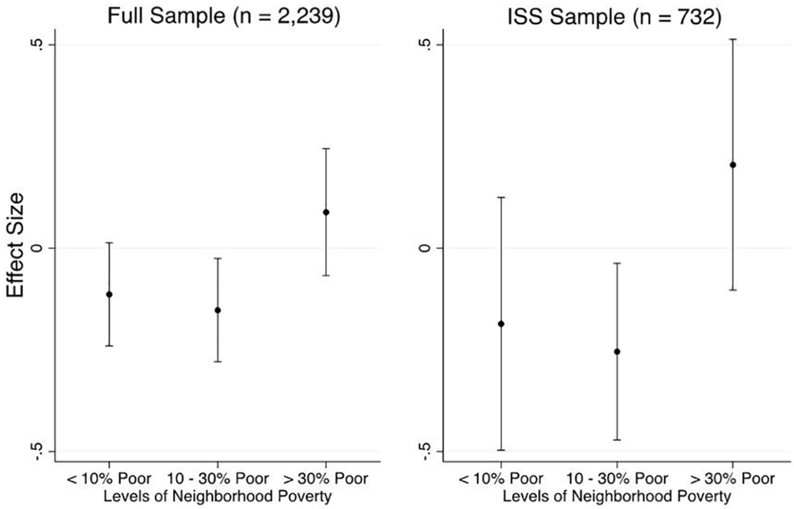

Figure 3. **ITT estimates of the effect of VPK on third-grade math achievement across neighborhood poverty.**

Note. ITT = intent to treat; VPK = Voluntary Prekindergarten program; ISS = Intensive Substudy.

this study finds that the effect of VPK differed across levels of neighborhood poverty, with the largest effects of VPK enrollment on third-grade reading achievement observed for children living in high-poverty neighborhoods. In contrast to what was observed for the full sample, however, there is no evidence that VPK enrollment or assignment affected the third-grade achievement of children living in low-poverty neighborhoods.

Intensive Substudy Sample: Reading Achievement

Columns (1) and (3) in Table 4 provide no evidence of a significant main effect of either VPK assignment or enrollment on third-grade reading achievement in the ISS sample. However, there exists considerable impact heterogeneity across levels of neighborhood poverty. As shown in Column (2) of Table 4, the estimated effect of VPK *assignment* on third-grade reading achievement was 0.67 *SD*s larger for children living at baseline in high- compared with low-poverty neighborhoods (β = 22.31, p = .016). Similarly, Column (4) shows that the estimated effect of VPK *enrollment* on third-grade reading achievement was 1.55 *SD*s larger for children living in high- compared with low-poverty neighborhoods (β = 51.45, p = .040).

As before, to help provide some intuition for the magnitude and practical significance of these interactions, the panels on the right side of Figures 1

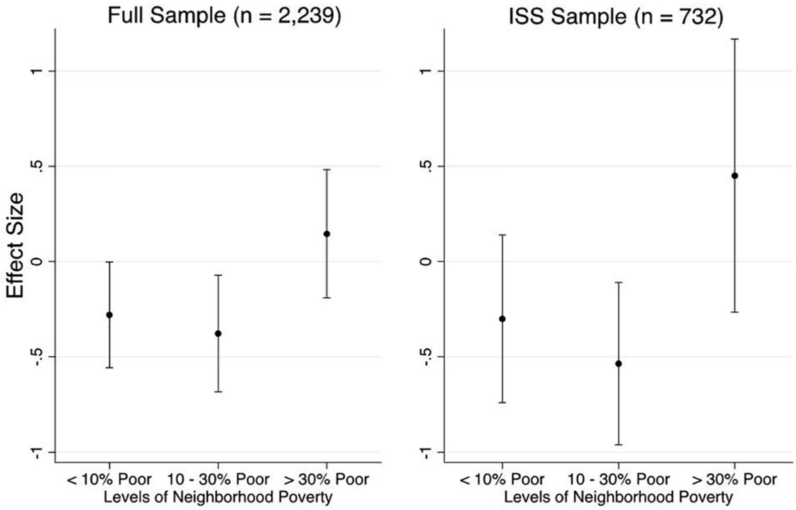

Figure 4. **LATE estimates of the effect of VPK on third-grade math achievement across neighborhood poverty.**

Note. LATE = local average treatment effect; VPK = Voluntary Prekindergarten program; ISS = Intensive Substudy.

and 2 show, for the ISS sample, the estimated effect of VPK assignment and enrollment, respectively, on third-grade reading achievement within each level of neighborhood poverty. These figures show that, among children living in high-poverty neighborhoods, VPK *assignment* was associated with a 0.55 *SD* gain in reading achievement in third grade (p = .023), and VPK *enrollment* was associated with a 1.36 *SD* gain (p = .065). As noted in the previous paragraph, no evidence was found that VPK assignment or enrollment affected the reading achievement of children in low-poverty neighborhoods. Neither was evidence found that VPK enrollment affected reading achievement of children in moderate-poverty neighborhoods.

Intensive Substudy Sample: Math Achievement

Columns (5) through (8) in Table 4 report ITT and LATE estimates of the main and moderated effect of VPK on third-grade math achievement in the ISS sample. Despite that the magnitude of the point estimates for the main effects and interaction terms were largely consistent with those observed for the full sample (i.e., negative main effects but potential heterogeneity across neighborhood types), only the main effect of VPK *enrollment* in the ISS sample was significant (β = −11.50, p = .085, effect size = −0.32 *SD*s);

the remainder of estimates were imprecisely estimated. Consequently, this study cannot rule out the possibility that, for the ISS sample, the overall effect of VPK *assignment* on math achievement is indistinguishable from zero and that the null effect of VPK assignment—as well as the effect of VPK enrollment—are invariant to the level of disadvantage in children's residential neighborhood at baseline.

Robustness Checks

Part B of the online supplement provides a series of additional analyses that show that substantive conclusions are robust to different thresholds of neighborhood poverty, alternative weighting strategies that account differential selection into the analytic samples, and the inclusion in the analytic sample of children who were unable to be geolocated to a valid street address at baseline but who could be geolocated to the centroid of the zip code in which they lived. Interestingly, no evidence is found that VPK effects varied across levels of neighborhood poverty when neighborhood poverty was measured at the census-tract- or zip-code-level as opposed to the block-group level. This finding is likely attributable to census tracts and zip codes covering larger geographic areas than block groups. For instance, the median size of block groups in Tennessee is roughly two square miles, whereas the median sizes of census tracts and zip codes in Tennessee are roughly 8 and 70 square miles, respectively. In other words, when conceptualizing neighborhoods at a broader geographic scale than the surrounding block groups in which children lived, this study found no evidence that neighborhood poverty moderates preK effects. That is, neighborhoods mattered at a granular level (Chetty, Friedman, Hendren, Jones, & Porter, 2018) such that incorporating characteristics of communities only miles away diminishes the moderating effects of community-level poverty rates. This finding indicates that whatever mechanisms might be driving the observed impact heterogeneity documented in the preceding pages likely operated at an acutely local scale.

Rival Moderator Analysis

Supplementary Appendix Tables I.1 to I.4 report results from the rival moderator analysis. This analysis examined the extent to which differential effects of VPK across levels of neighborhood poverty were explained by alternative childcare options, contextual risk factors, family SES, and urbanicity.[4] Overall, no evidence was found that the moderating role of neighborhood poverty on VPK effects was explained by lower-SES families being more likely to live in high-poverty neighborhoods. In fact, allowing the effect of VPK to vary across family SES amplified effect heterogeneity across levels of neighborhood poverty for both math and reading achievement. That is, impact heterogeneity across neighborhood poverty was simply not detecting differences in family SES, suggesting that there were likely "neighborhood effects" at play (Sharkey & Faber, 2014). For reading achievement, neighborhood effects operated to some extent through the distribution of alternative childcare options, which explained around 3% of the differential VPK effect across levels of neighborhood poverty, and contextual risk factors, which explained 6% to 7% of this variation. For math achievement, the relatively smaller differential effects described in the previous section were explained, in part, by contextual risk factors, which explained 1% to 3% of the differential VPK effects across levels of neighborhood poverty. Finally, urbanicity accounted for just 1% of the variation in VPK effects across levels of neighborhood poverty for reading achievement and 2% to 5% for math.

Discussion

The promise of preK finds its origins in evaluations of several small-scale early education demonstration programs, designed and run by researchers during the 1970s and 80s, that served low-income children living in disadvantaged areas of their respective cities. However, the expansion of prekindergarten in recent years to communities both within and outside metropolitan areas, to disinvested inner cities and affluent suburbs alike, raises the natural question of whether it is reasonable to expect similar effects for children living in such varied environments, environments that existing evidence suggests likely influence the counterfactual condition to which preK effects are compared (e.g., Wodtke & Harding, 2012). This study builds on these observations and leverages data from a recently conducted randomized controlled trial of VPK in Tennessee and demonstrates that VPK effects were contingent on the neighborhoods in which children lived.

This study estimates that the effect of VPK *assignment* on third-grade reading achievement for the full RCT sample was 0.39 *SD*s larger for children living in the high- compared with low-poverty neighborhoods in Tennessee, and that the average effect of VPK assignment for those living in high-poverty neighborhoods was positive and statistically significant. Similarly, this study estimates that the effect of VPK *enrollment* on third-grade reading achievement in the full RCT sample was fully 0.83 *SD*s larger for children in high- compared with low-poverty neighborhoods; likewise, the marginal effect of VPK enrollment on reading for those in high-poverty neighborhoods was significant and substantively meaningful. Evidence for impact heterogeneity with regard to math achievement was also observed, but this heterogeneity did not correspond with statistically meaningful effects for children living in high-poverty neighborhoods as it did for reading achievement. These conclusions were largely corroborated in the ISS sample, which allowed for a broader set of baseline controls including indicators of family SES and baseline levels of academic achievement.

Moreover, this study found no evidence that this impact heterogeneity across neighborhoods was simply detecting differences in family-level SES.

However, this study found evidence that the differential VPK effect across neighborhood contexts may have been attributable, in part, to the distribution of alternative childcare options and the types of risk factors children were exposed to in their neighborhood contexts. In particular, larger VPK effects were observed for children living in higher poverty neighborhoods, in part, because these children had fewer alternative childcare options and because they were overexposed to contextual risk factors relative to their peers living in low-poverty neighborhoods.

Last, despite finding compelling evidence that VPK effects differed across neighborhood contexts for reading achievement, less evidence of impact heterogeneity was observed for math achievement. Moreover, despite finding strong evidence of positive VPK effects on the reading achievement of children in high-poverty neighborhoods, no such evidence was found with regard to children's math achievement. This differential pattern of effects for reading versus math achievement could be the result of several factors. One speculative explanation for the persistence of effects for reading but not for math may be related to the practical emphasis of VPK curriculum. Despite curricular mandates established by the Tennessee Department of Education that required preschool teachers to cover both math and literacy skills, there is often variation in how preschool teachers implement curricular requirements (Hamre et al., 2010), with some evidence that instructional time for literacy can exceed that for math (Farran, Meador, Christopher, Nesbitt, & Bilbrey, 2017). If preschools in the analytic sample placed heavier emphasis on the development of early literacy skills than numeracy skills, then it is reasonable that effects on literacy, relative to those on math, would be more likely to make up for any literacy-related disadvantages at home and in communities.

A related explanation for why neighborhood poverty had a strong moderating effect on VPK with regard to reading achievement but less so for math achievement is that neighborhood poverty may not be as consequential for children's math relative to their reading achievement. In fact, Tables 3 and 4, which report moderated ITT and LATE estimates for the full and ISS samples, respectively, reveal that children living in a high-poverty neighborhood at baseline in the absence of VPK consistently performed worse, on average, than children living in lower-poverty neighborhoods on third-grade reading achievement across the full and ISS samples and with respect to ITT and LATE estimates. In contrast, the adjusted difference in third-grade math achievement between children living in high- versus low-poverty neighborhoods in the absence of VPK were of a smaller magnitude and were measured less precisely than those for third-grade reading achievement. Therefore, the differential moderating effects of neighborhood poverty on third-grade reading versus math achievement may simply be due to the fact that neighborhood poverty was more consequential for children's reading than math achievement.

These findings highlight several key considerations for theory, policy, and research for early childhood education. First, this study provides considerable evidence that understanding the distribution and prevalence of neighborhood poverty may be important for understanding nonschooling determinants of preK effect persistence (Bailey, Duncan, Odgers, & Yu, 2017). Moreover, that VPK effects persisted through third grade for children living in high-poverty neighborhoods is particularly noteworthy in light of several recent experimental and quasi-experimental studies that documented within their respective study populations near-universal fadeout of preK effects by third grade (Hill, Gormley, & Adelstein, 2015; Lipsey et al., 2017; Puma, Bell, Cook, & Heid, 2010; Puma et al., 2012). Second, that positive VPK impacts accrued for children in high-poverty neighborhoods is consistent with recent studies that document the importance of early childhood interventions for safeguarding children from periods of elevated developmental vulnerability due, in this case, to disadvantaged neighborhood contexts (Dodge et al., 2008).

Third, that effects hinged in either direction depending on children's level of neighborhood poverty suggest that a responsive policy agenda that aims to maximize a population-level preschool effect would likely need to include provisions that aim to increase the share of children from high-poverty neighborhoods in preschool classrooms. This inclusion could be accomplished in at least two ways. First, site-selection policies could prioritize establishing VPK centers in high-poverty areas, which would presumably expand VPK access to the children most likely to benefit. This is especially relevant because there is evidence in the current study that more children attended VPK in low-poverty neighborhoods than there were VPK children residing in low-poverty neighborhoods (see Supplementary Appendix Table A.1). This pattern of results suggests that there may have been inequities in VPK access based on neighborhood poverty. However, it is possible that suboptimal effects may be observed if increasing the number of preK classrooms in high-poverty neighborhoods concentrates large numbers of disadvantaged children in the same preK classroom as opposed to more economically integrated classrooms (Miller, Votruba-Drzal, McQuiggan, & Shaw, 2017; van Ewijk & Sleegers, 2010).

A second, alternative approach could focus on mitigating potential barriers, such as transportation and information gaps, that prevent children living in high-poverty neighborhoods from attending VPK centers in low-poverty neighborhoods (Greenberg, Michie, & Adams, 2018; J. F. Harding & Paulsell, 2018). Moreover, a policy agenda that endeavors to provide such access—that is, access that might increase the number of low-income children attending VPK classrooms in lower-poverty neighborhoods—would likely promote racial and socioeconomic integration of VPK classrooms as well given the strong correlation between neighborhood poverty, family SES, and racial composition (Reardon & Bischoff, 2011).

Limitations

Although this study extends prior early childhood research by highlighting the importance of residential contexts for understanding who did and did not benefit from VPK, this study is not without several limitations. The analytic samples suffered from nonnegligible attrition due to misreporting of baseline addresses and incomplete outcome data.[5] Moreover, the analytic samples were composed exclusively of economically disadvantaged children, whom TN-VPK targets for enrollment. Therefore, it is unclear whether the variation in treatment effects across neighborhood environments that was documented in the preceding pages would hold in a more socioeconomically diverse sample of children.[6] Finally, this study concerned academic outcomes only during third grade. Future research will need to identify whether the results described in this study pertain to preschool programs in other states, whether the moderated effects also pertain to behavioral or noncognitive outcomes (e.g., disciplinary outcomes, grade retention, attendance, executive function), and whether differential effects on outcomes of interest pertain to earlier grades or persist beyond third grade. Finally, in light of the limited explanatory power of the rival moderators examined in this study, future research will need to examine further the underlying mechanisms linking neighborhood poverty to children's engagement and response to early childhood programs.

Conclusion

Limitations notwithstanding, the current study provides important new evidence about variation in preK effects. Based on data from an RCT of a statewide preK program in Tennessee, this study finds that preK effects increase with level of poverty in children's residential neighborhoods, with positive effects accruing to low-income children in high-poverty neighborhoods but negative effects accruing to low-income children in low-poverty neighborhoods. This pattern was partially attributable to the distribution of alternative childcare options and the extent to which children were exposed to contextual risk factors in their neighborhoods. The analyses presented here should encourage future research into how children's neighborhoods relate to the implementation and efficacy of preschool interventions, especially those implemented at scale. In particular, if this study is to serve as any guide, recent efforts to expand preK to broader populations of children should double down on efforts at identifying *where* effects might be maximized.

Notes

The author would like to thank Dale Farran, Mark Lipsey, Kelley Durkin, Caroline Christopher, Georgine Pion, Mark Lachowicz, Matt Springer, and Walker Swain for their invaluable support and critical feedback as this paper developed. The author is also greatly indebted to the irreplaceable team of research assistants and project coordinators associated with the Tennessee Voluntary PreKindergarten Study, especially Janie Hughart, Richard Feldser, and Ilknur Sekmen, as well as to the Tennessee Department of Education and the Tennessee Education Research Alliance. The author would also like to thank Greg Duncan and participants in the 2017 Society for Research on Educational Effectiveness Annual Conference for a constructive discussion about this study. Additionally, the author wishes to thank Christy McGuire, Elizabeth Votruba-Drzal, Portia Miller, Tom Dee, Sean Reardon, and several anonymous reviewers who provided helpful comments about various aspects of this study. Financial support from the National Institutes of Health (NIH) is gratefully acknowledged (5R01HD079461-02).

Supplemental material for this article is available in the online version of the journal.

[1]Despite the TN-VPK Program being a statewide program with curricular and staffing mandates, the quality of VPK programs may vary across the state (Valentino, 2018). For instance, McCoy et al. (2015) found neighborhood poverty to be negatively correlated with the quality of Head Start classrooms. Therefore, children living in disadvantaged neighborhoods may also be more likely to attend a lower quality VPK program. Thus, if one was interested in the effect of statewide VPK controlling for VPK quality, then one would expect the estimated effects of VPK on the achievement of children in high-poverty neighborhoods to be plausible lower bounds.

[2]Using moderate-poverty neighborhoods as the referent group does not change substantive conclusions. In particular, virtually all instances in which preK effects differed between low- and high-poverty neighborhoods also varied between moderate- and high-poverty neighborhoods. The key implication is that positive preK effects were restricted to children in high-poverty neighborhoods, and these effects differed from those for children in lower-poverty neighborhoods.

[3]Differential access to alternative childcare options was measured by the density of childcare establishments per capita in the surrounding community. The number of childcare establishments was gathered from the 2010 Zip Business Patterns database and rescaled as a per capita rate based on the number of children younger than 5 years in the zip code. (Note that these estimates were only available at the zip code level.) Differential exposure to contextual risk factors was approximated with two measures: the share of single-parent families in children's neighborhood and the per capita crime rate in the surrounding community. The share of single-parent families was measured at the block-group level and gathered from the 2008–2012 American Community Survey. Per capita crime rate was gathered from the Tennessee Bureau of Investigation and measured as the number of violent and property crime arrests per capita in year 2010. (Similar to data about alternative childcare options, crime data were only available at the zip code level.) Family SES was measured by the mother's level of education. As described above, data about children's family-level SES were only available for the ISS sample. Finally, urbanicity was measured as a binary indicator of whether a child lived in an urbanized area or urban cluster, as defined by the U.S. Census Bureau. Urban areas form the cores of metropolitan statistical areas and contain 50,000 or more people, whereas urban clusters form the cores of micropolitan statistical areas and contain 2,500 to 50,000 people.

[4]Supplementary Appendix C (Figure C.1, Table C.1) shows that each rival moderator is correlated with neighborhood poverty, with children living in high-poverty neighborhoods having (1) less access to alternative childcare options, (2) overexposure to crime and violence, (3) overexposure to single-parent households in their communities, (4) mothers with less education, and (5) increased likelihood of living in an urban area. These correlations suggest that a rival moderator analysis—one that examines changes in the moderating capacity of neighborhood poverty after accounting for its correlation with each factor—is warranted.

[5]This study showed that results were robust to the exclusion of sampling weights that accounted for students' differential probability of being included in the analytic samples based on observed characteristics (see Supplementary Tables C.8 to C.12). Moreover, results were robust to the inclusion in the analytic sample of a subset of children who could not be geolocated to a street address but could be geolocated to the centroid of their home zip code. Thus, there is little reason to suspect that analytic samples were substantively different from their full RCT and ISS samples.

[6]Prior research has shown that children in nonpoor families are less vulnerable to adverse neighborhood conditions (Wodtke & Harding, 2012), suggesting that the low-income children included in the sample may have been especially susceptible to the type of impact heterogeneity observed. Consequently, findings should not be generalized to universal preK programs that "target" a broader range of children.

ORCID iD

Francis A. Pearman II https://orcid.org/0000-0002-9516-8302

References

Bailey, D., Duncan, G. J., Odgers, C. L., & Yu, W. (2017). Persistence and fadeout in the impacts of child and adolescent interventions. *Journal of Research on Educational Effectiveness*, *10*, 1934–5747. doi:10.1080/19345747.2016.1232459

Berruate-Clement, J. R. (1984). *Changed lives: The effects of the Perry Preschool Program on youths through age 19* [Monographs of the High/Scope Educational Research Foundation, Number Eight]. Retrieved from http://files.eric.ed.gov/fulltext/ED313128.pdf

Bloom, H. S., & Weiland, C. (2015). *Quantifying variation in Head Start effects on young children's cognitive and socio-emotional skills using data from the National Head Start Impact Study*. Retrieved from https://www.mdrc.org/sites/default/files/quantifying_variation_in_head_start.pdf

Brooks-Gunn, J., Johnson, A. D., & Leventhal, T. (2010). Disorder, turbulence, and resources in children's homes and neighborhoods. In *Chaos and its influence on children's development: An ecological perspective* (pp. 155–170). Washington, DC: American Psychological Association. doi:10.1037/12057-000

Bureau of the Census. (2008). Census Block Group Program for the 2010 Decennial Census—Final criteria. *Federal Register*, *73*(51). Retrieved from https://www2.census.gov/geo/pdfs/reference/fedreg/bg_criteria.pdf

Chetty, R., Friedman, J. N., Hendren, N., Jones, M. R., & Porter, S. R. (2018, October). *The Opportunity Atlas: Mapping the childhood roots of social mobility* (NBER Working Paper No. 25147). Cambridge, MA: National Bureau of Economic Research. doi:10.3386/w25147

Chetty, R., Hendren, N., & Katz, L. F. (2015). The effects of exposure to better neighborhoods on children: New evidence from the Moving to Opportunity experiment. *American Economic Review*, *106*(4), 1–88. doi:10.1257/aer.20150572

Cohen, S. (2017). *3 principles to improve outcomes for children and families*. Cambridge, MA: Harvard University, Center on the Developing Child. Retrieved from https://46y5eh11fhgw3ve3ytpwxt9r-wpengine.netdna-ssl.com/wp-content/uploads/2017/10/HCDC_3PrinciplesPolicyPractice.pdf

Cooper, B. R., & Lanza, S. T. (2014). Who benefits most from Head Start? Using latent class moderation to examine differential treatment effects. *Child Development*, *85*, 2317–2338. doi:10.1111/cdev.12278

Dodge, K. A., Greenberg, M. T., & Malone, P. S. (2008). Testing an idealized dynamic cascade model of the development of serious violence in adolescence. *Child Development*, *79*, 1907–1927. doi:10.1111/j.1467-8624.2008.01233.x

Farran, D. C., Meador, D., Christopher, C., Nesbitt, K. T., & Bilbrey, L. E. (2017). Data-driven improvement in prekindergarten classrooms: Report from a partnership in an urban district. *Child Development*, *88*, 1466–1479. doi:10.1111/cdev.12906

Fitzpatrick, M. D. (2008). Starting school at four: The effect of universal pre-kindergarten on children's academic achievement. *B. E. Journal of Economic Analysis & Policy*, *8*(1), 1–38. doi:10.2202/1935-1682.1897

Flora, C. B., Flora, J. L., & Gasteyer, S. P. (2015). *Rural communities: Legacy+ change*. Boulder, CO: Westview Press.

Galster, G. (2012). The mechanism(s) of neighbourhood effects: Theory, evidence, and policy implication. In M. van Ham, D. Manley, N. Bailey, L. Simpson, & D. Maclennan (Eds.), *Neighbourhood effects research: New perspectives* (pp. 23–56). New York, NY: Springer.

Gormley, W. T., Phillips, D., & Gayer, T. (2008). Preschool programs can boost school readiness. *Science*, *320*, 1723–1724. doi:10.1126/science.1156019

Green, T. (2015). Places of inequality, places of possibility: Mapping "opportunity in geography" across urban school-communities. *Urban Review*, *47*, 717–741. doi:10.1007/s11256-015-0331-z

Greenberg, E., Michie, M., & Adams, G. (2018). *Expanding preschool access for children of immigrants*. Washington, DC: Urban Institute. Retrieved from https://www.urban.org/sites/default/files/publication/96546/expanding_preschool_access_for_children_of_immigrants_0.pdf

Hamre, B. K., Justice, L. M., Pianta, R. C., Kilday, C., Sweeney, B., Downer, J. T., & Leach, A. (2010). Implementation fidelity of MyTeachingPartner literacy and language activities: Association with preschoolers' language and literacy growth. *Early Childhood Research Quarterly*, *25*, 329–347. doi:10.1016/J.ECRESQ.2009.07.002

Harding, D. J. (2003). Counterfactual models of neighborhood effects: The effect of neighborhood poverty on dropping out and teenage pregnancy. *American Journal of Sociology*, *109*, 676–719. doi:10.1086/379217

Harding, J. F., & Paulsell, D. (2018). *Improving access to early care and education: An equity-focused policy research agenda*. Princeton, NJ: Mathematica.

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

Jacob, B. A., & Wilder, T. (2010). *Educational expectations and attainment* (NBER Working Paper No. 15683). Cambridge, MA: National Bureau of Economic Research. Retrieved from https://www.nber.org/papers/w15683

Jargowsky, P. (1997). *Poverty and place: Ghettos, barrios, and the American city*. New York, NY: Russell Sage Foundation.

Jocson, K. M., & Thorne-Wallington, E. (2013). Mapping literacy-rich environments: Geospatial perspectives on literacy and education. *Teachers College Record*, *115*(June), 1–24.

Karch, A. (2010). Policy feedback and preschool funding in the American states. *Policy Studies Journal*, *38*, 217–234. doi:10.1111/j.1541-0072.2010.00359.x

Lipsey, M. W., Farran, D., & Durkin, K. (2017). Effects of the Tennessee Prekindergarten Program on children's achievement and behavior through third grade. *Early Childhood Research Quarterly*, *45*(4th quarter 2018), 155–176. doi:10.1016/j.ecresq.2018.03.005

Lipsey, M. W., Hofer, K. G., Dong, N., Farran, D. C., & Bilbrey, C. (2013, August). *Evaluation of the Tennessee Voluntary Prekindergarten Program: Kindergarten and first grade follow-up results from the randomized control design*. Nashville, TN: Peabody Research Institute. Retrieved from https://my.vanderbilt.edu/tnprekevaluation/files/2013/10/August2013_PRI_Kand1stFollowup_TN-VPK_RCT_ProjectResults_FullReport1.pdf

McCoy, D. C., Connors, M. C., Morris, P. A., Yoshikawa, H., & Friedman-Krauss, A. H. (2015). Neighborhood economic disadvantage and children's cognitive and social-emotional development: Exploring Head Start classroom quality as a mediating mechanism. *Early Childhood Research Quarterly*, *32*, 150–159. doi: 10.1016/j.ecresq.2015.04.003

McCoy, D. C., Morris, P. A., Connors, M. C., Gomez, C. J., & Yoshikawa, H. (2016). Differential effectiveness of Head Start in urban and rural communities. *Journal of Applied Developmental Psychology*, *43*, 29–42. doi:10.1016/j.appdev.2015.12.007

Miller, P., Votruba-Drzal, E., McQuiggan, M., & Shaw, A. (2017). Pre-K classroom-economic composition and children's early academic development. *Journal of Educational Psychology*, *109*, 149–165. doi:10.1037/edu0000137

Milner, H. R. (2015). *Rac(e)ing to class: Confronting poverty and race in schools and classrooms*. Cambridge, MA: Harvard Education Press.

Morris, P. A., Connors, M., Friedman-Krauss, A., McCoy, D., Weiland, C., Feller, A., & Page, L. (2017). New findings on impact variation from the Head Start Impact Study: Informing the scale-up of early childhood programs. *AERA Open*, *4*(2), 1–16. doi:10.1177/2332858418769287

Puma, M., Bell, S., Cook, R., & Heid, C. (2010). *Head Start impact study: Final report*. Washington, DC: U.S. Department of Health and Human Services. Retrieved from http://files.eric.ed.gov/fulltext/ED507845.pdf

Puma, M., Bell, S., Cook, R., Heid, C., Broene, P., Jenkins, F., . . . Downer, J. (2012). *Third grade follow-up to the Head Start Impact Study: Final report*. Washington, DC: U.S. Department of Health and Human Services. Retrieved from http://files.eric.ed.gov/fulltext/ED539264.pdf

Reardon, S. F., & Bischoff, K. (2011). Income inequality and income segregation. *American Journal of Sociology*, *116*, 1092–1153. doi:10.1086/657114

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, *70*, 41–55.

Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. New York, NY: Wiley.

Schochet, P. Z., & Deke, J. (2014). *Understanding variation in treatment effects in education impact evaluations: An overview of quantitative methods*. Washington, DC: U.S. Department of Education. Retrieved from https://ies.ed.gov/ncee/pubs/20144017/pdf/20144017.pdf

Sharkey, P., & Faber, J. W. (2014). Where, when, why, and for whom do residential contexts matter? Moving away from the dichotomous understanding of neighborhood effects. *Annual Review of Sociology, 40*, 559–579. doi:10.1146/annurev soc-071913-043350

Sharkey, P., Tirado-Strayer, N., Papachristos, A. V., & Raver, C. C. (2012). The effect of local violence on children's attention and impulse control. *American Journal of Public Health*, *102*, 2287–2293. doi:10.2105/ajph.2012.300789

Small, M. L. (2008). Four reasons to abandon the idea of the ghetto. *City and Community*, *7*, 389–398. doi:10.1111/j.1540-6040.2008.00271_8.x

Swenson, K. (2008). *Child care arrangements in urban and rural areas*. Washington, DC: Office of the Assistant Secretary for Planning and Evaluation.

Tate, W. F. (2008). "Geography of opportunity": Poverty, place, and educational outcomes. *Educational Researcher*, *37*, 397–411. doi:10.3102/0013189X08326409

Tennessee Department of Education. (2017). *Scope of services: Voluntary Pre-K 2017-2018*. Retrieved from https://www.ecschools.net/docs/district/welc/vpk%20scope%20of%20services%2017-18.pdf?id=661

Valentino, R. (2018). Will public pre-K really close achievement gaps? Gaps in prekindergarten quality between students and across states. *American Educational Research Journal*, *55*, 79–116. doi:10.3102/0002831217732000

van Ewijk, R., & Sleegers, P. J. C. (2010). The effect of peer socioeconomic status on student achievement: A meta-analysis. *Educational Research Review*, *5*, 134–150.

VanderWeele, T. J., & Knol, M. J. (2011). Interpretation of subgroup analyses in randomized trials: heterogeneity versus secondary interventions. *Annals of Internal Medicine*, *154*, 680–683. doi:10.7326/0003-4819-154-10-201105170-00008

Votruba-Drzal, E., Miller, P., & Coley, R. L. (2016). Poverty, urbanicity, and children's development of early academic skills. *Child Development Perspectives*, *10*, 3–9. doi:10.1111/cdep.12152

Wilson, W. J. (2012). *The truly disadvantaged: The inner city, the underclass, and public policy*. Chicago, IL: University of Chicago Press.

Wodtke, G. T. (2013). Duration and timing of exposure to neighborhood poverty and the risk of adolescent parenthood. *Demography*, *50*, 1765–1788. doi:10.1007/s13524-013-0219-z

Wodtke, G. T., Elwert, F., & Harding, D. J. (2016). Neighborhood effect heterogeneity by family income and developmental period. *American Journal of Sociology*, *121*, 1168–1222. doi:10.1086/684137

Wodtke, G. T., & Harding, D. J. (2012). *Poor families, poor neighborhoods: How family poverty intensifies the impact of* concentrated disadvantage (Population Studies Center Research Report No. 12-776). Retrieved from https://www.psc.isr.umich.edu/pubs/abs/7739

Wodtke, G. T., Harding, D. J., & Elwert, F. (2011). Neighborhood effects in temporal perspective: The impact of long-term exposure to concentrated disadvantage on high school graduation. *American Sociological Review*, *76*, 713–736. doi:10.1177/0003122411420816

Yoshikawa, H., Weiland, C., Brooks-Gunn, J., Burchinal, M., Gormley, W., Ludwig, J., & Zaslow, M. J. (2013). *Investing in our future: The evidence base on preschool education*. New York, NY: Foundation for Child Development.

Manuscript received November 9, 2017
Final revision received July 16, 2019
Accepted August 2, 2019