Annenberg Institute
BROWN UNIVERSITY

EdWorkingPaper No. 21-494

The Long-Term Effect of North Carolina Pre-Kindergarten is Larger in School Districts with Lower Rates of Growth in Academic Achievement

Robert C. Carr
Duke University

Tyler W. Watts
Teachers College, Columbia University

Jade M. Jenkins
University of California, Irvine

Yu Bai
Duke University

Ellen Peisner-Feinberg
The University of North Carolina at Chapel Hill

Clara G. Muschkin
Duke University

Helen F. Ladd
Duke University

Kenneth A. Dodge
Duke University

Prior research has found that public investments in North Carolina's pre-kindergarten program—NC Pre-K—generated positive effects on student reading and math achievement through eighth grade (Bai et al., 2020). This study examined whether the effect of NC Pre-K funding exposure is moderated by the educational environments children subsequently experience during elementary and middle school. The NC Pre-K effect on student reading and math achievement in eighth grade was found to be larger in school districts with lower rates of growth in academic achievement. These findings suggest that public investments in early childhood education may be particularly beneficial in the long term for children who subsequently experience low-growth school environments—consistent with a dynamic substitutability hypothesis of combined effects.

VERSION: October 2024

Suggested citation: Carr, Robert C., Tyler Watts, Jade M. Jenkins, Yu Bai, Ellen S. Peisner-Feinberg, Clara G. Muschkin, Helen F. Ladd, and Kenneth A. Dodge. (2024). The Long-Term Effect of North Carolina Pre-Kindergarten is Larger in School Districts with Lower Rates of Growth in Academic Achievement. (EdWorkingPaper: 21-494). Retrieved from Annenberg Institute at Brown University: https://doi.org/10.26300/jrgc-8n94

The Long-Term Effect of North Carolina Pre-Kindergarten is Larger in School Districts with Lower Rates of Growth in Academic Achievement

Robert C. Carr [a]

Tyler W. Watts [b]

Jade M. Jenkins [c]

Yu Bai [a]

Ellen S. Peisner-Feinberg [d]

Clara G. Muschkin [a]

Helen F. Ladd [a]

and

Kenneth A. Dodge [a]

[a] Duke University
[b] Teachers College, Columbia University
[c] University of California, Irvine
[d] The University of North Carolina at Chapel Hill

Author Note

Corresponding author: Robert C. Carr, Duke University, 302 Towerview Road, Durham, NC 27708-0539. robert.carr@duke.edu.

Funding: This research study was supported by grant funding from the Eunice Kennedy Shriver National Institute of Child Health and Human Development (1R01HD095930-01A1).

Declaration of interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this manuscript.

Acknowledgements: The authors wish to thank the North Carolina State Center for Health Statistics, the North Carolina Department of Public Instruction, the North Carolina Division of Child Development and Early Education, the North Carolina Partnership for Children, and the North Carolina Education Research Data Center for providing the relevant data files and guidance. Thank you to the members of the Consortium on Early Childhood Intervention Impact (CECII) for their helpful feedback on earlier versions of this work.

Abstract

Prior research has found that public investments in North Carolina's pre-kindergarten program—NC Pre-K—generated positive effects on student reading and math achievement through eighth grade (Bai et al., 2020). This study examined whether the effect of NC Pre-K funding exposure is moderated by the educational environments children subsequently experience during elementary and middle school. The NC Pre-K effect on student reading and math achievement in eighth grade was found to be *larger* in school districts with *lower* rates of growth in academic achievement. These findings suggest that public investments in early childhood education may be particularly beneficial in the long term for children who subsequently experience low-growth school environments—consistent with a dynamic substitutability hypothesis of combined effects.

Keywords: Early childhood education; long-term effects; reading; mathematics.

Introduction

Public investments in early childhood education (ECE) can result in substantial benefits to society, especially if those investments target high-quality programming for children whose families experience economic insecurity (Heckman, 2006). Numerous studies have investigated the rollout of ECE programs across communities and over time (see Cascio, 2021 for a review). These studies have documented many positive effects across a variety of developmental, educational, and life outcomes. For example, a series of quasi-experimental studies investigated the rollout of funding for North Carolina's statewide pre-kindergarten (pre-K) program—NC Pre-K. Findings document the positive effect of NC Pre-K funding exposure on students' standardized test scores in reading and math—with effects evident as early as third grade (Dodge et al., 2016; Ladd et al., 2013) and persistent through eighth grade (Bai et al., 2020). However, previous findings also demonstrate that the NC Pre-K effect can vary across the educational environments children experience after preschool (Watts et al., 2023). Still, there remains some uncertainty regarding the mechanisms by which school-age educational environments may enhance, diminish, or add to the long-term effects of ECE. Understanding how school environments influence the persistence of preschool effects is a topic of interest to both policy and academic audiences alike (Bailey et al., 2017; Phillips et al., 2017).

The current study examines two key dimensions of the educational environment children experience during elementary and middle school which may influence the persistence of preschool effects. We examine these dimensions as moderators of the NC Pre-K funding effect. Following prior research by Reardon (2019), we operationalize two dimensions of the school-age educational environment in North Carolina public schools: (1) *average achievement* is the average level of academic achievement among third-grade students in the school district and (2) *achievement growth* is the average rate of growth in academic achievement among students in the school district as they progress from third to eighth grade (measured as the linear slope). Watts et al. (2023) have found that the positive effect of NC Pre-K funding exposure on fifth grade achievement is *larger* in elementary schools where average achievement is *lower*. Average achievement may reflect a broad range of educational inputs available to students in their schools and communities (e.g., neighborhood characteristics, average family resources in a community), whereas academic achievement growth may better reflect the educational inputs uniquely related to public education (Reardon, 2019). The current study examines the moderating effects of average achievement *and* academic achievement growth in relation to the effect of NC Pre-K exposure on student achievement outcomes in eighth grade.

In the current study, we found that school-district average achievement and achievement growth were weakly correlated—suggesting that these measures reflect distinct dimensions of the educational environment in North Carolina public school districts. We also found independent positive effects of NC Pre-K funding exposure, school-district average achievement, and school-district achievement growth in relation to student reading and math achievement in eighth grade. The effect of NC Pre-K funding exposure was not reliably moderated by school-district average achievement. However, we found that the positive effect of NC Pre-K funding exposure was moderated by school-district achievement growth. Specifically, the magnitude of the NC Pre-K effect on student reading and math outcomes was larger in school districts where the rate of academic achievement growth was lower. The results of this study offer new insights into the ways in which school-age school environments contribute to the persistence of ECE effects in the long term.[1]

North Carolina's Pre-K Program

NC Pre-K is a state-funded educational program designed to enhance children's early learning and social-emotional skills prior to kindergarten entry. When NC Pre-K was established in 2001, North Carolina was among the latter half of states to adopt a public pre-K program (Cohen-Vogel et al., 2020).[2] The program grew quickly during the following years. By 2010, 24% of the state's 4-year-old population was participating in the program (*N* = 31,197) (Barnett et al., 2010). That proportion has remained relatively consistent over time (e.g., 24% served in 2019) (Friedman-Krauss et al., 2020). The North Carolina program is targeted to low-income 4-year-olds (plus children with disabilities), so the cap on participation is well below the full population.

The NC Pre-K program has maintained high quality standards since its inception. For example, during its first year of operation, the NC Pre-K program met 7 of 10 quality benchmarks established by the National Institute of Early Education Research (e.g., class size and staff-child ratios, teacher qualifications and training, learning standards, health and nutrition services, and program monitoring) (Barnett et al., 2003). Just four years later, the NC Pre-K program had increased its standards to meet all 10 quality benchmarks (Barnett et al., 2006) and continues to maintain these standards (Friedman-Krauss et al., 2024). Local NC Pre-K providers are required to meet a variety of program standards designed to ensure a high-quality educational experience for all children. For example, all NC Pre-K providers are required to maintain a four- or five-star license—the highest quality ratings under North Carolina's child care licensing system. Moreover, NC Pre-K has sought to promote uniformity in program standards across the state through quality monitoring and continuous quality improvement initiatives.

The NC Pre-K program targets four-year-old children who meet certain eligibility criteria, including a gross family income at or below 75% of the state median income, or if the child has at least one of the following factors: educational or developmental delay, an identified disability, a chronic health condition, or limited English proficiency (Barnett et al., 2010).[3] NC Pre-K does not create stand-alone pre-K classrooms; instead, the state provides a NC Pre-K funded "slot" to qualifying children in a variety of classroom-based settings that meet the state quality standards, including public schools, Head Start, and private child care centers (both for-profit and nonprofit). Therefore, a key feature of the NC Pre-K program concerns the blended composition of NC Pre-K classrooms, which include children who receive an NC Pre-K funded slot *as well as* children who do not receive a funded slot. Through this strategy of funding slots for individual children in blended classroom settings, the program seeks to promote high-quality preschool for funded children as well as all other children enrolled in the same classrooms. During our study period, estimates suggest that two-thirds of the children enrolled in a participating NC Pre-K classroom received a funded slot, while one-third of children in participating classrooms were not directly funded by the NC Pre-K program (Peisner-Feinberg et al., 2006).[4] Therefore, the benefits of the NC Pre-K funding could spill over to children who were not directly funded by the program but enrolled in the same classroom—suggesting the possibility of indirect effects of NC Pre-K. A study by Muschkin et al. (2024) found some evidence of indirect effects from the expansion of NC Pre-K. In that study, having a higher percent of former NC Pre-K students enrolled in the elementary school was associated with positive outcomes among elementary school teachers—including teacher satisfaction with their working environment, as well as teacher and principal retention from one year to the next.

In the current study, we leveraged the differential rollout of funding for NC Pre-K across North Carolina's 100 counties during an 18-year period (described in greater detail below). We used variability in the amount of funding that was allocated to communities over time in order to identify the community-wide effect of NC Pre-K. With this approach, we examined the effect of variations in NC Pre-K funding exposure across counties and years for the population of students in North Carolina.

ECE Exposure

A robust body of research has investigated the effects of ECE by examining the rollout of ECE programs operating at a large scale (see Cascio, 2021 for a review). Rollout can refer to establishing, implementing, and/or scaling up ECE programs. These studies provide a wealth of evidence to indicate that exposure to ECE can have positive effects for children across a variety of developmental, educational, and life outcomes. This includes evidence from two randomized experiments in which the timing of ECE rollout was randomized across rural communities in Cambodia (Bouguen et al., 2018) and Mozambique (Martinez et al., 2017). Beyond these two randomized experiments in Africa, previous research has relied on quasi-experimental studies where ECE rollout was not randomized but varied naturally (and seemingly randomly). This includes studies of ECE programs in several European countries, in Canada, as well as in the United States.

In the U.S., positive effects of ECE have been documented in quasi-experimental studies of state-funded pre-K program rollout in North Carolina (Bai et al., 2020; Dodge et al., 2016; Ladd et al., 2013; Muschkin et al., 2015), Georgia (Cascio & Schanzenbach, 2013; Fitzpatrick, 2008), Oklahoma (Cascio & Schanzenbach, 2013), and South Carolina (Williams, 2019), as well as the rollout of the federal Head Start preschool program (Bailey et al., 2021; Barr & Gibbs, 2022; Johnson & Jackson, 2019; Kose, 2021; Thompson, 2018). All of these quasi-experimental studies in the U.S. have documented the positive effects of ECE on educational outcomes—often focusing on longer-term outcomes, such as educational attainment in adulthood (e.g., Bailey et al., 2021). A subset of these U.S. studies has examined academic achievement outcomes based on standardized tests scores, and those studies also document positive effects (Bai et al., 2020; Cascio & Schanzenbach, 2013; Dodge et al., 2016; Fitzpatrick, 2008; Ladd et al., 2013; Thompson, 2018; Williams, 2019). Moreover, the U.S. studies that have examined academic achievement outcomes focused on medium- to long-term outcomes in elementary school or middle school. These studies have not yet provided findings related to the immediate, short-term effect of ECE rollout on children's academic readiness for school.

Studies of ECE program rollout typically compare children *with* and *without* the possibility of exposure to ECE programming *before* and *after* the rollout of an ECE program. Moreover, these studies measure population-wide effects that could be conferred through direct benefits to ECE program participants as well as spillover effects to non-participants. Many studies of ECE program rollout have utilized applications of the quasi-experimental difference-in-differences (DiD) design—including all prior studies conducted in the U.S. The conventional DiD approach relies on a two-by-two comparison. The first difference is the change over time within one area *before* and *after* the implementation of ECE. The second difference concerns change over the same timeperiod in a comparison area where ECE is not implemented. It is the difference between these two differences that the conventional DiD approach relies on. Many DiD studies examine the staggered rollout of ECE at different times (or for different cohorts) across many different areas that implement ECE programs to varying degrees. In these staggered rollout studies, the ECE effect represents a weighted average of all possible two-by-two comparisons that can be constructed from the data (Goodman-Bacon, 2021). In studies where the ECE rollout is both staggered and measured as a continuous treatment (e.g., NC Pre-K funding exposure is measured as a continuous treatment in the current study), the ECE effect represents a weighted average of all possible two-by-two comparisons among units that experienced different levels (or doses) of the continuous treatment (Callaway et al., 2024).

A series of DiD studies conducted by our research team has examined the staggered rollout of North Carolina's NC Pre-K program (Bai et al., 2020; Dodge et al., 2016; Ladd et al., 2013; Muschkin et al., 2015). These studies leveraged quasi-experimental variation in the allocation of funding for NC Pre-K across North Carolina's 100 counties—including *between*-county variation as well as *within*-county variation over time. These studies have documented the positive effect of NC Pre-K funding exposure on student academic achievement outcomes in reading and math, as well as reductions to grade retention and special education placement. These effects were evident in third grade and continued to persist through eighth grade.

Findings from ECE program rollout studies can be considered alongside studies that investigate ECE program participation—including randomized control trials that randomly assign children to an offer of ECE program participation as well as quasi-experimental studies that compare ECE program participants to non-participants. This includes three random-assignment studies of ECE programs (Lipsey et al., 2018; Peisner-Feinberg et al., 2019; Puma et al., 2010), as well as numerous quasi-experimental studies (e.g., propensity score matching or regression discontinuity) comparing children who participated in ECE programs to non-participating peers. The results of ECE program participation studies are reviewed elsewhere in greater detail (Burchinal et al., 2024; Cascio, 2021; Phillips et al., 2017; Yoshikawa et al., 2013). In short, these studies have documented robust evidence of positive, short-term effects of ECE participation on children's academic readiness to begin elementary school. However, the findings are mixed in studies that follow children into elementary school and beyond—especially with regard to the effects on academic achievement outcomes. In these studies, there is evidence of diminishing effects during elementary school as well as evidence of positive effects that persist through elementary school and into later grades. A study of Tennessee's pre-K program has also documented the emergence of adverse effects during elementary school, after finding initially positive short-term effects (Lipsey et al., 2018). The effect of participation in North Carolina's NC Pre-K program has been investigated in a random-assignment study (Peisner-Feinberg et al., 2020; Peisner-Feinberg et al., 2019) as well as several quasi-experimental studies (Burchinal et al., 2022; Carr et al., 2024; Carr et al., 2021; Peisner-Feinberg & Schaaf, 2010, 2011). Similar to the broader literature, these studies document robust evidence of the positive, short-term effect of NC Pre-K participation. However, evidence concerning the medium-term effect of NC Pre-K participation is mixed—including some evidence of positive as well as null effects during the early elementary grades. The longer-term effect of NC Pre-K participation has not yet been investigated.

In summary, a wealth of research has investigated the effects of ECE programs operating at scale. A series of studies conducted by our research team has examined the rollout of funding for North Carolina's NC Pre-K program using an application of the DiD design. Our findings document positive, long-term, and population-wide effects resulting from the rollout of funding for NC Pre-K. In the current study, we build on our previous work to examine how the effect of NC Pre-K funding exposure is moderated by the educational environments children experience after they graduate from preschool and progress through school.

ECE and Subsequent Educational Environments During School

There may be certain conditions under which the positive effects of ECE are more or less likely to persist in the long run. It seems likely that persistence of the ECE effect will depend on the subsequent educational environments children experience as they progress through school. However, understanding how the effects of subsequent educational environments interact with ECE effects has not been resolved. There are two leading hypotheses regarding the interaction between the effects of ECE and subsequent educational environments: *dynamic complementarity* and *dynamic substitutability* (Bailey, Duncan, et al., 2020). Both hypotheses can be evaluated by a test of moderation (or statistical interaction) between measures of ECE and the subsequent educational environment.

The first hypothesis—dynamic complementarity—suggests that the long-term effects of ECE will be *larger* in the context of *higher*-quality educational environments during school (Bailey et al., 2017; Cunha et al., 2006; Phillips et al., 2017). This hypothesis is rooted in economic theory of human capital formation which suggests that early investments in children's skill development are not productive in the long run unless they are followed up by later investments across the lifespan (Cunha et al., 2006). In the education research literature, dynamic complementarity is also called the *sustaining environments* hypothesis (see Bailey et al., 2017). According to this hypothesis, the short-term effects of ECE will carry over into long-term effects when a higher-quality K-12 education enables children to build on the skills they gained through ECE while acquiring more advanced skills during school.

An alternative hypothesis—dynamic substitutability—suggests that the long-term effects of ECE will be *larger* in the context of *lower*-quality educational environments during school and *smaller* in *higher*-quality educational environments (Bailey, Duncan, et al., 2020). Dynamic substitutability suggests that the ECE effect will compensate for lower-quality educational environments during school (Abenavoli, 2019; Bailey, Duncan, et al., 2020). For example, Abenavoli (2019) describes how high-quality ECE programming may promote resilience among children who subsequently enroll in lower-quality educational environments. When examining the interaction between educational inputs across preschool and school, the dynamic substitutability hypothesis suggests that children with ECE program exposure will be better off relative to their peers without ECE exposure in lower-quality school environments (i.e., a positive effect of ECE in lower-quality school environments). Reciprocally, higher-quality school environments will boost the skill development of children who missed out on ECE and enable them to catch up to their peers with previous ECE exposure (Abenavoli, 2019; Bailey, Duncan, et al., 2020).

From a statistical standpoint, support for the dynamic complementarity hypothesis would be evidenced by positive main effects of ECE and the subsequent educational environment as well as a positive interaction between these two inputs; while support for the dynamic substitutability hypothesis would be evidenced by a negative interaction (Bailey, Jenkins, et al., 2020). There is also a third hypothesis—the *additivity* hypothesis—which suggests that ECE effects are *not* moderated by later educational environments, but still exert independent effects on the outcome of interest. Support for the additivity hypothesis would be evidenced by positive main effects of ECE and the subsequent educational environment along with a null or statistically non-significant interaction between these two inputs (see Watts et al., 2023).

Findings from existing research are largely mixed and do not provide a clear consensus regarding the dynamic effects of ECE and subsequent educational environments (see review by Bailey, Jenkins, et al., 2020). To date, we know of only two studies that have investigated the dynamic effects of ECE program rollout and subsequent educational environments. A study by Johnson and Jackson (2019) provides evidence of dynamic complementarity, while another study conducted by our research team provides evidence of dynamic substitutability (Watts et al., 2023). Both studies are described in greater detail below.

Johnson and Jackson (2019) examined the historical rollout of funding for the federal Head Start program between 1965 and 1980 as well as per pupil spending in K-12 education. The long-term educational outcome examined in this study pertained to high school graduation. The findings provide support for the dynamic complementarity hypothesis, whereby the positive effect of Head Start exposure was *larger* for children who experienced *higher* per pupil spending in K-12 education.

In contrast, a study by our research team examined the rollout of funding for North Carolina's NC Pre-K program in a modern policy context (between 1993 and 2010) as well as a variety of ecological environments children experience across development (Watts et al., 2023). That study found support for the dynamic substitutability hypothesis based on student academic achievement outcomes in fifth grade. A variety of moderating measures were examined in that study. School-wide average achievement emerged as a moderator of the NC Pre-K effect on student achievement outcomes in fifth grade. Specifically, the positive effect of NC Pre-K funding was *larger* in the context of elementary schools with lower-levels of school-wide average achievement and *smaller* in the context of elementary schools with higher-levels of school-wide average achievement.

It is possible, that dynamic effects may operate differently for different educational outcomes (i.e., high school graduation vs. fifth-grade academic achievement) and different program types and times (i.e., the historical rollout of Head Start vs. the rollout of a modern state-funded pre-K program). Additionally, the rollout of Head Start funding *and* per pupil spending were shown to represent plausibly exogenous and independent variation in the study by Johnson and Jackson (2019). The rollout of NC Pre-K funding was also shown to represent plausibly exogenous variation in the study by Watts et al. (2023), but could not be presumed for measures of the subsequent school environment in that study (including school-wide average achievement and per pupil spending).

The *educational* environments children experience during school likely influence the extent to which ECE effects will persist in the long run. However, previous studies by Johnson and Jackson (2019) and Watts et al. (2023) have examined moderating measures that indirectly assess the K-12 educational environment—including per pupil expenditures and average achievement. The current study builds on previous work by Watts et al. (2023) to examine the dynamic effect of *academic achievement growth* as well as *academic achievement growth* in relation to the effect of NC Pre-K funding on student achievement outcomes in eighth grade. If ECE continues to exert long-term effects on academic achievement outcomes throughout school, those effects are expected to differ based on the extent to which public schools promote growth in academic achievement among students.

Measuring Average Achievement and Achievement Growth in School Districts

Previous research by Reardon (2019) has utilized standardized test score data from students in public-school districts across the U.S. to operationalize two dimensions of the educational environment children experience during school. These measures include (1) the average level of academic achievement among third grade students in the school district (i.e., average achievement) and (2) the average rate of growth in academic achievement among students in the school district as they progress from third to eighth grade (i.e., achievement growth). In that study, the correlation between average achievement and achievement growth was weak (*r* = –0.13). Additionally, average achievement was highly correlated with a measure of community-wide socioeconomic status (*r* = 0.68), whereas achievement growth was only moderately correlated (*r* = 0.32). These correlations provide evidence of discriminant validity and suggest that average achievement and achievement growth capture distinct dimensions of the school-age educational environment—average achievement may reflect a broad range of educational inputs available to students in their schools and communities beyond the contribution of public education (e.g., neighborhood characteristics, average family resources in a community), whereas achievement growth may better capture the educational inputs uniquely related to public education (Reardon, 2019).

Building on the approach by Reardon (2019), we measured average achievement and achievement growth across public school districts in North Carolina. Educational theory and research suggest many ways in which school districts can directly and indirectly shape the educational experiences of students (Blazar & Schueler, 2022). For example, case studies demonstrate that school-district administrators make important decisions that directly influence the structural characteristics and culture of schools, including decisions about curriculum; decisions about hiring, staffing, and salaries of teachers and principals; as well as decisions about the professional development and evaluation of teachers and principals (Chenoweth, 2017, 2021). In extreme cases, school district administrators can fundamentally reform and/or reorganize schools through school closures, consolidations, and the implementation of accountability systems. The North Carolina public school system is one in which student achievement measures are commonly used for high-stakes accountability and performance monitoring. School-district administrators can also influence the culture of school environments.

For example, Chenoweth (2021) described effective school districts as having a culture in which the leadership cultivates the belief that all children are capable of success and that the adults are responsible for ensuring children succeed (p. 131). Indeed, a comprehensive case study of Chicago Public Schools identified *effective leaders* as the most critical ingredient in building a culture of school improvement (Bryk et al., 2010). In sum, school-district leadership can shape the structural organization and culture of schools, and, in turn, these features of the school environment can influence students' academic skill development (Eccles & Roeser, 2010). Finally, the current study's focus on school districts rather than schools as the unit of analysis for the educational environment was warranted because variation across the 115 public school districts in North Carolina closely corresponds to the level of analysis used to examine the rollout of NC Pre-K funding across North Carolina's 100 counties (15 counties have two districts; typically, a city and a county district).

The Current Study

The long-term effects of North Carolina's NC Pre-K program may depend on the subsequent educational environments that children experience during school. The extant evidence is mixed regarding how subsequent educational environments will play a role in moderating the long-term effects of ECE (Bailey, Jenkins, et al., 2020). The current study examined two dimensions of the educational environment children experience in North Carolina public school districts—average achievement and achievement growth—as moderators of the NC Pre-K effect on student reading and math achievement in eighth grade. We used a linked administrative dataset that includes 18 cohorts of more than one million children who were born in North Carolina and then went on to progress through eighth grade in North Carolina public schools.

The current study proceeded in two phases. In the first phase of our study, we measured average achievement and achievement growth in North Carolina public school districts using a growth curve modeling approach. Specifically, we measured (a) the average level of academic achievement among third-grade students in the school district and (b) the average rate of growth in academic achievement among students in the school district as they progressed from third to eighth grade. In the second phase of our study, we examined the effect of NC Pre-K funding exposure in relation to student reading and mathematics outcomes in eighth grade. We used a DiD design to examine variation in the rollout of NC Pre-K program funding allocated to North Carolina's 100 counties across 18 program years. We simultaneously tested for moderation of the NC Pre-K funding effect by the measures of average achievement and achievement growth in North Carolina public school districts.

We considered three hypotheses. Support for the *dynamic complementarity* hypothesis would be evidenced by positive effects of NC Pre-K program funding and school-district educational environment as well as a positive interaction between these inputs. Support for the *dynamic substitutability* hypothesis would be evidenced by positive effects of these inputs as well as a negative interaction between inputs. Finally, support for the *additivity* hypothesis would be evidenced by positive effects of these inputs, but no reliable interaction between inputs.

Methods

Data Sources

Our study used data on the population of students enrolled in North Carolina public schools in third grade through eighth grade between the 1996 and 2019 school years (school records obtained from the North Carolina Department of Public Instruction [NCDPI]) as well as a subpopulation of these students who were also born in North Carolina between January 1, 1987, and August 31, 2005 (birth records obtained from North Carolina State Center for Health Statistics).[5] Data on the full population of students were used in the Phase I analyses while data on the subpopulation of students were used in the Phase II analyses (described in further detail in the *Analyses* section below). To combine these data, the identified birth records and school records of individual students were matched by the North Carolina Education Research Data Center (NCERDC) at Duke University. In the current study, 74% of all birth records were matched to a school record. We also utilized data on state allocations for NC Pre-K program funding obtained from the NCDPI, and publicly available data on each of North Carolina's 100 counties obtained from the North Carolina Office of State Budget and Management (NCOSBM).

Measures

NC Pre-K funding exposure

Data on NC Pre-K program funding were obtained from the state administrative agency, which allocated funds to program contractors in local communities (i.e., counties) on a state fiscal year basis.[6] We converted all funds into 2019-dollar values based on the Consumer Price Index. To calculate county-level exposure to NC Pre-K funding, we divided the amount of NC Pre-K funds awarded to each county during each fiscal year by the number of age-eligible children (i.e., age 4) in each county during each fiscal year (based on county-level population estimates from NCOSBM):

$$NC\ Pre\ K\ funding\ exposure = \frac{NC\ Pre\ K\ funds\ in\ FY}{4\ year\ olds\ in\ FY}$$

Children were assigned to the value of county-level NC Pre-K funding exposure during the fiscal year in which they were 4 years old and eligible to benefit from the program. The NC Pre-K program was established during state fiscal year 2002. After the program was established, the average county-level exposure to NC Pre-K funding grew steadily and reached a peak of

$2,030 in 2010, with considerable variation within and across counties over the study period (see Figure S1). The average county-level exposure to NC Pre-K funding between program years 2002-2010 was $1,064. We estimated that the program directly served 28% of all four-year-old children in North Carolina in 2010 (the last year of program funding considered in the current study).

Student Achievement in Reading and Mathematics

Standardized assessments of reading and mathematics achievement were administered to students in North Carolina public schools at the end of each grade between third grade and eighth grade. Our study used scores from these end-of-grade (EOG) tests administered during all school years between 1996 and 2019. The EOG tests are state-mandated, standardized assessments with well-documented psychometric characteristics that satisfy the requirements of the U.S. Department of Education (Bazemore & Van Dyk, 2004; Mbella et al., 2016a, 2016b; North Carolina Department of Public Instruction, 2009; Sanford, 1996). Separate editions of the EOG tests were administered across the range of years included in our panel study (i.e., four editions of the EOG Reading and five editions of the EOG Math). To account for changes in the scaling of EOG test editions across years, we converted the EOG scores into *Z*-scores (across all editions) by subject, school year, and grade level, resulting in a *Z*-score of 0.00 equal to the statewide average for students in each grade during each school year.

Covariates

Student covariates. Data on individual students were obtained from (a) the birth records data and (b) the school records data. The student covariates include student sex (Female = 0, Male = 1; school records); student economic disadvantage (No = 0, Yes = 1; school records); student birth weight categorized as extremely low, very low, low, normal (reference category), or high (birth records); student race categorized as non-Hispanic White (reference category), African American, Hispanic, Native American, Asian, or mixed race (school records); the quarter of the calendar year in which the child was born, categorized as quarter 1, 2, 3, and 4 (with quarter 3 as the reference category; adjusted for the kindergarten entry date and defined as the youngest children eligible for kindergarten entry during the school year; birth records); maternal education in years (birth records); parent marital status (Not married = 0, Married = 1; birth records); maternal age in years (birth records); no dad information on birth record (birth records); mother immigrant (No = 0, Yes = 1; birth records); student was first born (No = 0, Yes = 1; birth records); mother race categorized as non-Hispanic White (reference category), African American, Hispanic, Native American, Asian, or mixed race (birth records).

School-district covariates. Data on North Carolina's public school districts across years were obtained from the NCERDC. The school-district covariates include the per pupil expenditures from local, state, and federal funding sources (converted to 2019 $s), the percent of economically disadvantaged students in the school district, and the total number of students in the school district's membership during eighth grade.[7]

County covariates. Socio-demographic data on each of North Carolina's 100 counties across years were obtained from (a) the NCOSBM or (b) the birth records data aggregated from student-level information to county-level information. These county covariates include the percent of births to African American mothers, Hispanic mothers, and low education mothers (< 12 years of education) (birth records); the percent of population receiving Food Stamps and the percent of population receiving Medicaid (NCOSBM); the total number of births (log-transformed; birth records); the total population (log transformed; NCOSBM); and median family income (converted to 2019 $s; NCOSBM).

In the current study, we also controlled for children's exposure to funding for North Carolina's Smart Start program. The Smart Start program was established in 1993 by the North Carolina General Assembly to serve children (and their families) from birth to age 4 in three main areas: child care subsidy and quality improvement, family support services, and child health. The North Carolina Partnership for Children awards funds to local Smart Start partnerships to provide program services in each of North Carolina's 100 counties. Local decision-making is a hallmark of the Smart Start program, as the activities of local partnerships in each of North Carolina's 100 counties are guided by a board of community members who select the services that best meet the needs of children and families in their community (North Carolina Partnership for Children, 2019). The decentralized nature of the Smart Start program leads to varied implementation of program services across counties and over time. Therefore, the implemented program can best be described as a pool of financial resources with guidelines for spending. Data on Smart Start program funding were obtained from the North Carolina Partnership for Children. Smart Start funds were allocated to program contractors/counties on a state fiscal year basis (converted to 2019 $s).[8] Similar to our measure of NC Pre-K funding exposure, we calculated county-level exposure to Smart Start funding in two steps. First, we divided the amount of Smart Start funds awarded to each county during each fiscal year by the number of age-eligible children (i.e., ages birth–4) in each county during each fiscal year (based on county-level population estimates from NCOSBM).[9] We then summed the annual Smart Start funding exposure in each county across each five-year fiscal year period in which children in our panel were eligible to benefit from the program (i.e., between birth–age 1, 1–2, 2–3, 3–4, and 4–age 5). Children were assigned to the value of county-level Smart Start funding exposure during the five-year fiscal year period in which they were eligible to benefit from the program.

Analyses

Analyses proceeded in two phases. In the first phase, we measured school-district average achievement in third grade and achievement growth from third to eighth grade among students in North Carolina public school districts. We then examined correlates of average achievement and achievement growth. In the second phase of analysis, we examined the effect of NC Pre-K program funding exposure on student achievement outcomes in eighth grade, and we tested whether this effect was moderated by school-district average achievement and achievement growth. All analyses were completed in SAS® version 9.4.

Phase I: Measuring School-District Average Achievement and Achievement Growth

In Phase I of our analyses, we measured school-district average achievement and achievement growth for 18 cohorts of students who progressed from third to eighth grade in each of North Carolina's 115 public school districts (see Figure S3).[10, 11, 12] We began with third grade because it was the lowest grade level at which standardized achievement tests were administered. We used the reading and math scores drawn from the full population of students enrolled in North Carolina public schools with valid EOG test score data.[13] The growth cohorts were defined as all students who took the third grade through eighth grade reading and/or math assessments during a six-year period.[14, 15] For example, the first growth cohort was defined by all students who took the third-grade reading and/or math assessment in 1996, the fourth-grade assessment in 1997, the fifth-grade assessment in 1998, the sixth-grade assessment in 1999, the seventh-grade assessment in 2000, and/or the eighth-grade assessment in 2001.

We estimated a multi-level growth curve model for each of the 18 growth cohorts, with separate models estimated for reading and math scores ($18 \times 2 = 36$ models total). All models were estimated using the HPMIXED procedure in SAS® version 9.4. The generalized equation for the multi-level growth curve models to measure school-district average achievement and achievement growth is shown below (with separate models calculated for each of the 19 cohorts):

Level-1 (time; *t*):

$$Y_{tij} = \pi_{0ij} + \pi_{1ij} time_{tij} + e_{tij}$$

Level-2 (students; *i*):

$$\pi_{0ij} = \beta_{00j} + r_{0ij}$$
$$\pi_{1ij} = \beta_{10j} + r_{1ij}$$

Level-3 (school districts; *j*):

$$\beta_{00j} = \gamma_{000} + u_{00j}$$
$$\beta_{10j} = \gamma_{100} + u_{10j}$$

Variance Components:

$$e_{tij} \sim N(0, \sigma_{tij})$$
$$r_{0ij} \sim N(0, \sigma_{0ij})$$
$$r_{1ij} \sim N(0, \sigma_{1ij})$$
$$u_{00j} \sim N(0, \sigma_{00j})$$
$$u_{10j} \sim N(0, \sigma_{10j})$$

$\boldsymbol{Y_{tij}}$ is the test score at time *t* (grade) for student *i* in school district *j*; $time_{tij}$ is coded as 0 at third grade, 1 at fourth grade, 2 at fifth grade, 3 at sixth grade, 4 at seventh grade, and 5 at eighth grade (separate models were calculated for reading and mathematics scores)

$\boldsymbol{\pi_{0ij}}$ is the expected achievement level (i.e., intercept) for student *i* in school district *j* in third grade

$\boldsymbol{\pi_{1ij}}$ is the expected achievement growth (i.e., linear slope) for student *i* in school district *j* during each grade between third and eighth grade

$\boldsymbol{\beta_{00j}}$ is the expected achievement level (i.e., intercept) for students in school district *j* in third grade

$\boldsymbol{\beta_{10j}}$ is the expected achievement growth (i.e., linear slope) for students in school district *j* during each grade between third and eighth grade

In this model, repeated assessments of students' reading or mathematics skills over time (*t*; Level-1) were nested within students (*i*; Level-2), and students were nested within school districts (*j*; Level-3). Time was coded as 0 at third grade, 1 at fourth grade, 2 at fifth grade, 3 at sixth grade, 4 at seventh grade, and 5 at eighth grade. Test scores (Y_{tij}) were modeled as a function of (1) an intercept term centered at the third-grade assessment score to represent the expected achievement level for student *i* in school district *j* in third grade (π_{0ij}) and the expected achievement level for students in school district *j* in third grade (β_{00j}) as well as (2) a linear slope term to represent the expected achievement growth for student *i* in school district *j* during each grade between third and eighth grade (π_{1ij}) and the expected achievement growth for students in school district *j* during each grade between third and eighth grade (β_{10j}). The intercept term was allowed to vary randomly between students (r_{0ij}) and school districts (u_{00j}), and the slope term was also allowed to vary randomly between students (r_{1ij}) and school districts (u_{10j}). All of the variance terms were parameterized to be normally distributed random variables with a mean of zero. We specified an unstructured covariance matrix to allow the random intercepts and slopes and their variances to be correlated with one another at Level-2 and at Level-3 in order to allow for a systematic relation between intercepts and slopes. We used restricted maximum likelihood (REML) to estimate the variance components and the "residual" method to calculate denominator degrees of freedom. Students with at least one test score were included in these analyses and maximum likelihood was used to account for missing data (Singer & Willett, 2003). No covariates were included in these analyses in order to calculate unconditional estimates of the intercepts and slopes.

Based on the results of this model, we derived estimates of average achievement (i.e., intercept; β_{00j}) and achievement growth (i.e., slope; β_{10j}) in reading/math for each of the 115 school-districts across the 18 cohorts for use in the subsequent analyses. Specifically, we derived the empirical best linear unbiased predictions (EBLUPs) for the realizations of the random intercept, slope, and nested errors. This model was conceptually similar to model 8.15–8.17 described by Bryk and Raudenbush (1992). Based on the results of these analyses, we examined correlations between school-district economic disadvantage, average achievement, and achievement growth in order to better understand the discriminant validity of these measures.

Phase II: Eighth-Grade Student Achievement Analyses

In Phase II of our analyses, we examined the effect of NC Pre-K funding exposure on student reading and math achievement outcomes in eighth grade. NC Pre-K funding exposure varied across North Carolina's 100 counties during an 18-year period. We operationalized a DiD design using linear regression analyses with a two-way fixed effect (TWFE) specification—including county and year fixed effects. In these TWFE analyses, the effect of NC Pre-K funding exposure represents the weighted average of between-county funding effects and within-county funding effects across time—weighted by variance and sample size (Kropko & Kubinec, 2020). A generalized equation is shown below:

$$O_{iscByPySy} = \beta_0 + \beta_1 NCPK\ funding_{cPy} + \boldsymbol{X}_i + \boldsymbol{Z}_{sSy} + \boldsymbol{C}_{cBy} + \alpha_{cBy} + \gamma_{iPy} + \alpha * \gamma_{icByPy} + \epsilon_{iscByPySy}$$

In this model, O is the eighth-grade achievement score (with separate models estimated for reading and mathematics scores) for student *i*, enrolled in eighth grade in school district *s*, born in county *c*, born in birth year *By*, exposed to program funding during program year *Py*, and enrolled in eighth grade during school year *Sy*;[16] β_1 is the main effect of NC Pre-K funding exposure which is scaled such that a 1-unit change is equal to a change from $0 to $2,030 (i.e., during our study period, the average county-level exposure to NC Pre-K funding reached a peak of $2,030 in 2010; see Figure S1); X is a vector of effects for the student-level covariates; Z is a vector of effects for the school-district-level covariates; C is a vector of effects for the county-level covariates; α is the fixed effect for county of residence to account for within-county variation in program funding over time; γ is the fixed effect for program year to account for between-county variation in program funding; $\alpha * \gamma$ is the linear time trend for each county across program years to account for trends in student achievement outcomes among students born within counties over time (see Dynarski et al., 2018 for more details); and ϵ is the residual term. The program funding, county-level covariates, and county fixed effect were aligned to students' county of residence and year of birth.[17] The school district covariates were aligned to the students' school district and school year during eighth grade. The dependent variables and continuous independent variables were grand mean standardized for the analysis sample with *M* = 0, *SD* = 1 (with separate standardization conducted for the reading and math outcome analysis samples). Dichotomous independent variables were centered, but not standardized (with separate centering conducted for the reading and math outcome analysis samples). All models were estimated with robust standard errors clustered at the county level.

In a two-way fixed effect analysis with a continuous treatment (e.g., NC Pre-K funding exposure), a key assumption concerns the linearity of the treatment effect estimate. In particular, if the treatment effect is non-linear, then the two-way fixed effect estimate may differ meaningfully from the average causal response parameter (Callaway et al., 2024). To provide a test of treatment effect homogeneity, we extend the previous model by adding a quadratic term for NC Pre-K funding exposure using the generalized equation shown below:

$$O_{iscByPySy} = \beta_0 + \beta_1 NCPK\ funding_{cPy} + \beta_2 NCPK\ funding_{cPy}^2 + \boldsymbol{X}_i + \boldsymbol{Z}_{sSy} + \boldsymbol{C}_{cBy} + \alpha_{cBy} + \gamma_{iPy} + \alpha * \gamma_{icByPy} + \epsilon_{iscByPySy}$$

In this model, we tested for evidence of non-linearity by estimating a quadratic (i.e., non-linear) effect of NC Pre-K funding exposure on student eighth-grade achievement outcomes (β_2)

Assuming a linear effect of NC Pre-K funding exposure, we extended the generalized equation to examine moderation of the NC Pre-K effect by school-district average achievement and achievement growth. The generalized equation to examine moderation of the NC Pre-K effect is shown below:

$$O_{iscByPySy} = \beta_0 + \beta_1 NCPK\ funding_{cPy} + \beta_2 Avg\ Ach_{sSy-1} + \beta_3 Ach\ Growth_{sSy-1}$$
$$+\beta_4 NCPK\ funding\ x\ Avg\ Ach_{scPySy-1} + \beta_5 NCPK\ funding\ x\ Ach\ Growth_{scPySy-1}$$
$$+ \boldsymbol{X}_i + \boldsymbol{Z}_{sSy} + \boldsymbol{C}_{cBy} + \alpha_{cBy} + \gamma_{iPy} + \alpha * \gamma_{icByPy} + \epsilon_{iscByPySy}$$

In this model, β_2 estimates the main effect of school-district average achievement (i.e., average achievement in reading for the reading outcome model and average achievement in math for the math outcome model); β_3 estimates the main effect of school-district achievement growth (i.e., achievement growth in reading for the reading outcome model and achievement growth in math for the math outcome model); β_4 estimates the interaction between NC Pre-K funding exposure and school-district average achievement to assess whether the relationship between NC Pre-K funding and the outcome changes across the different level of school-district average achievement; and β_5 estimates the interaction between NC Pre-K funding exposure and school-district achievement growth to assess whether the relationship between NC Pre-K funding and the outcome changes across the different level of school-district achievement growth. The school-district average achievement and achievement growth measures were grand mean standardized for the analysis sample with *M* = 0, *SD* = 1 (with separate standardization conducted for the reading and math outcome analysis samples). As a result, the main effect of NC Pre-K funding exposure can be interpreted in the context of the interaction terms (i.e., the effect of NC Pre-K funding exposure at the average level of school-district average achievement and achievement growth). The school-district average achievement and achievement growth measures were aligned to the students' school district during eighth grade, and we implemented a "current year minus 1" approach to assign values to students in a given school year on the basis of values derived from the prior school year, with a different set of students (e.g., students enrolled in eighth grade in 2002 were assigned school-district average achievement and achievement growth values from cohort #1, which was comprised of students enrolled in eighth grade in 2001, and so on).[18] This approach was similar to that used by Andrabi et al. (2011).

All models were estimated using the SURVEYREG procedure in SAS® version 9.4. These analyses relied on the subsample of students in North Carolina public schools with birth records matched to school records. From this matched subsample, only 4% and 7% of students had missing data on one or more study variables for the reading and math outcome analyses, respectively, and we applied listwise deletion to these cases.

Results

Tables 1 and S3 summarize the descriptive information on all study variables at the student level.

Phase I: Measuring School-District Average Achievement and Achievement Growth

We measured school-district average achievement and achievement growth in reading and mathematics for 18 cohorts of students who progressed from third to eighth grade between 1996 and 2018. Based on EOG scores transformed to *Z*-scores, we found that the value of school-district average achievement in third grade was –0.13 (*SD* = 0.22) for reading and –0.14 (*SD* = 0.23) for math (see Table S1). These findings indicated that the average third grade achievement of students in the public-school districts in our sample was slightly more than one tenth of a standard deviation below the statewide average.[19] Based on EOG scores transformed to *Z*-scores, we found that school-district achievement growth from third to eighth grade was –0.02 points (*SD* = 0.20) for reading and –0.01 points (*SD* = 0.20) for math.[20] These findings indicated that, on average, students in the public-school districts in our sample showed only slight declines in their statewide rank-order standing in *Z*-scores between third and eighth grade.

We examined correlates of school-district average achievement and achievement growth (see Table S2). First, we found that school-district average achievement and achievement growth were weakly correlated for both reading (r = 0.04, p = .04) and math (r = –0.16, p < .001). These findings suggest that average achievement and achievement growth may reflect distinct dimensions of the educational environments children experience in North Carolina public school districts. Second, we found that average achievement in reading was highly positively correlated with average achievement in math (r = 0.92, p < .001) and achievement growth in reading was highly correlated with achievement growth in math (r = 0.66, p < .001). These findings offer some evidence of convergent validity for each measure of the educational environment. Third, we found that school-district economic disadvantage was highly negatively correlated with average achievement in reading (r = –0.65, p < .001) and math (r = –0.65, p < .001). Alternatively, we found that school-district economic disadvantage showed a weak negative correlation with achievement growth in reading (r = –0.19, p < .001) and math (r = –0.07, p = .002). Consistent with Reardon (2019), our findings suggest that average achievement may reflect a broad range of socioeconomic conditions of the community where the school district is located, while achievement growth may better reflect the unique contributions of public education toward promoting student achievement during the elementary and middle school grades.

Phase II: Eighth-Grade Student Achievement Analyses

Table 2 summarizes the results of regression analyses to examine the main effect of NC Pre-K funding exposure in relation to student reading and math achievement in eighth grade. We found a positive main effect of NC Pre-K funding exposure for reading (β = 0.08, $p < .001$) and math (β = 0.11, $p = .01$). In separate models, we examined the non-linear effect of NC Pre-K funding exposure by adding a quadratic term for funding to the primary specification. This model provides a test of treatment effect homogeneity across levels of NC Pre-K funding exposure. The quadratic effect of NC Pre-K funding exposure was not statistically significant in relation to eighth-grade achievement in reading (β = 0.00, $p = .76$) and math (β = 0.03, $p = .37$). Moreover, the *R*-square value remained unchanged to the hundredth decimal place in models *with* and *without* the quadratic term. These findings provide evidence of treatment effect homogeneity and suggest that the relation between NC Pre-K funding exposure and student reading and math achievement was best modeled as a linear function across the range of funding exposure observed in our dataset. Therefore, in our subsequent analyses, we only examined the linear effect of NC Pre-K funding exposure.

Table 3 summarizes the results of regression analyses to examine moderation of NC Pre-K funding exposure by school-district average achievement and school-district achievement growth.[21] In relation to student reading achievement in eighth grade, we found positive main effects of NC Pre-K funding exposure (β = 0.07, $p = .003$), school-district average achievement in reading (β = 0.04, $p < .001$), and school-district achievement growth in reading (β = 0.03, $p < .001$). Similarly, in relation to student math achievement in eighth grade, we found positive main effects of NC Pre-K funding exposure (β = 0.09, $p = .03$), school-district average achievement in math (β = 0.07, $p < .001$), and school-district achievement growth in math (β = 0.08, $p < .001$). For both reading and math outcomes, the effect of NC Pre-K funding exposure was not reliably moderated by school-district average achievement in reading (β = –0.01, $p = .14$) nor average achievement in math (β = –0.01, $p = .53$). However, we found that the effect of NC Pre-K funding exposure was moderated by school-district achievement growth in reading (β = –0.02, $p < .001$) and achievement growth in math (β = –0.03, $p = .007$) such that the magnitude of the NC Pre-K effect decreased as achievement growth increased.

When we probed the interaction in relation to eighth-grade reading outcomes, we found that the effect of NC Pre-K funding exposure was positive at each level of achievement growth, but became smaller as achievement growth increased from *low* levels (e.g.., 1 standard deviation [*SD*] below the average; β = 0.09, $p < .001$) to *average* levels (β = 0.07, $p = 0.003$) and then *high* levels (e.g., 1 *SD* above the average; β = 0.04, $p = 0.07$) (see Figure 1). Similarly, when we probed the interaction in relation to eighth-grade math outcomes, we found that the effect of NC Pre-K funding exposure was positive at each level of achievement growth, but became smaller as achievement growth increased from *low* levels (β = 0.11, $p = .003$) to *average* levels (β = 0.09, $p = 0.03$) and then *high* levels (β = 0.06 $p = 0.15$). Therefore, the long-term effect of NC Pre-K funding was more robust for children in school districts where the rate of academic achievement growth was lower.

Discussion

High-quality ECE can have positive effects on student achievement outcomes. However, what remains unclear is how the educational environments children experience during school will moderate the long-term effects of ECE (Bailey, Jenkins, et al., 2020). In the current study, we examined the rollout of financial investments in North Carolina's NC Pre-K program. Specifically, we examined the effect of NC Pre-K funding exposure in relation to student reading and math achievement in eighth grade. Simultaneously, we examined moderation of the NC Pre-K funding effect by the academic achievement environments in North Carolina public school districts. We considered two hypotheses regarding the mechanisms by which subsequent educational environments might moderate the long-term effects of NC Pre-K—*dynamic complementarity* and *dynamic substitutability*.

We found support for a dynamic substitutability effect of exposure to NC Pre-K funding, such that increases in NC Pre-K funding exposure during preschool had a positive effect on student reading and math achievement in eighth grade, and this effect was *larger* in school districts with *lower* rates of growth in reading and math academic achievement. These findings suggest that the long-term effects of NC Pre-K are likely to be largest in the context of later educational environments that do not facilitate higher rates of growth in academic achievement. It is important to note that these findings revealed that students in higher-growth school districts were more likely to attain higher achievement in eighth grade compared to students in lower-growth school districts, regardless of their level of NC Pre-K funding exposure. Therefore, the developmental benefit of higher-growth school districts appeared to act as a substitute for children with low *or* no exposure to NC Pre-K funding during early childhood. Simultaneously, higher exposure to NC Pre-K funding appeared to buffer children against the consequences of subsequently enrolling in a lower-growth school district—the academic achievement of students in lower-growth school districts more closely approximated those of students in higher-growth school districts if they were also exposed to higher-levels of NC Pre-K program funding.

Dynamic substitutability could be evident for NC Pre-K because the program has long maintained a focus on promoting children's school readiness skills, including frequent and high-quality instruction related to early academic skills (e.g., Peisner-Feinberg et al., 2013). Hence, the developmental benefits resulting from higher exposure to NC Pre-K may have buffered students from lower rates of growth in academic achievement during school. Alternatively, in communities with low or no exposure to NC Pre-K, enrolling in a school district with higher rates of growth in academic achievement may have enabled those students to make up the gains in underlying reading and math skills that NC Pre-K could have otherwise conferred. In sum, it is possible that we observed dynamic substitutability because the benefits of a school readiness skill-building intervention closely corresponded to the benefits conferred through higher rates of academic skill development in high-growth school districts.

Our study focused on educational environments in school districts. We found that school-district average achievement and achievement growth both had positive main effects on student reading and math achievement in eighth grade. However, it is important to note that average achievement—but not achievement growth—showed a strong, negative correlation with the level of socioeconomic disadvantage among students in the school districts. This suggests that average achievement might measure a wide range of educational inputs available to students in their communities. Alternatively, achievement growth might better reflect the unique contributions of public education toward promoting student skill development. Therefore, these different dimensions of the educational environment during school can represent different pathways toward raising student achievement. For example, the benefits of enrolling in a school district with higher levels of average student achievement may begin to accrue as soon as children step inside the kindergarten door—impacting early skill attainment—while the benefits of achievement growth may require multiple years of exposure to accrue benefits—impacting later skill attainment. Additionally, the educational benefits of enrolling in a school district with higher levels of average achievement may extend beyond the experiences students have during the school day to their experiences in community centers and after school programs. A previous study by our research team found that NC Pre-K funding exposure had a positive effect on student achievement outcomes in fifth grade, and that effect was diminished in the context of elementary schools with lower-levels of school-wide average achievement (Watts et al., 2023). However, that study was not able to examine achievement growth in relation to student achievement outcomes. The current study points to achievement growth—rather than average achievement—as an important factor in differentiating the long-term benefits of NC Pre-K funding exposure.

Finally, we found main effects of NC Pre-K funding exposure in the current study that may be considered small by conventional standards (e.g., a difference of $2,030 in NC Pre-K funding exposure per 4-year-old child in the population led to a 0.08 *SD* unit increase in eighth grade reading achievement). However, these effects were educationally meaningful according to empirical benchmarks established for educational interventions (Kraft, 2020). These small magnitude effects can be further contextualized by the moderate cost of NC Pre-K funding exposure (i.e., an average of $2,030 per 4-year-old children in counties during the 2009-10 program year would considered moderate, and not small or large according to Kraft, 2020).[22] Overall, the reasonable cost of investments in NC Pre-K produced population-wide improvements in student achievement outcomes that remained educationally meaningful in the long-term—suggesting that even larger investments could produce greater impact in the future.

In the debate concerning the persistence (or not) of ECE program effects on student achievement, the findings of our research should encourage educational stakeholders and policymakers to consider how subsequent educational environments play a role in determining whether program effects are maintained or diminished in the long run. While our previous studies have documented NC Pre-K's enduring effects on academic achievement outcomes through elementary school and middle school (Bai et al., 2020; Dodge et al., 2016; Ladd et al., 2013), findings from the current study point out the possible conditional nature of these effects.

We found that the average rate of academic achievement growth in the school district made a difference in determining whether ECE program effects on student achievement were enhanced or diminished in the long run. Historically, exposure to NC Pre-K might have had the largest long-term effect for students who went on to enroll in school districts where growth in academic achievement was lowest. Today, efforts to expand access to NC Pre-K programming have become a policy priority in North Carolina (Carr & Peisner-Feinberg, 2021, September 13). If access to NC Pre-K programming becomes more readily available in the future, educators and policymakers might seek to promote greater continuity in high-quality educational environments across preschool and school. Several frameworks have been developed to promote educational alignment (e.g., Kauerz & Coffman, 2013). Consistent with these alignment frameworks, we advocate for a view of NC Pre-K programming and its benefits as part of a continuum of high-quality educational inputs to be built upon sequentially across development (Carr, 2021).

Study Limitations

The current study focused on educational environments in school districts as the unit of analysis, and our measures of average achievement and achievement growth were both positively associated with student achievement in eighth grade. Nonetheless, it is possible that educational environments may vary between schools within school districts. Examinations of school-level variation was outside the scope of our study because many schools did not include the full range of grade levels needed to measure average achievement and achievement growth from third to eighth grade. Additionally, we did not examine moderation of the NC Pre-K effect on student outcomes in third through seventh grade because school-district average achievement and achievement growth were measured between third and eighth grade.

Conclusion

Our study adds important nuance to a body of evidence documenting the positive, long-term effects of North Carolina's NC Pre-K program. While our prior research has shown that greater exposure to financial investments in NC Pre-K can have population-wide effects on student reading and math achievement that endure through the end of middle school—nine years after program exposure—the current study contributes new evidence regarding the role of school-age educational environments in differentiating these long-term effects. Specifically, we found that the positive effect of NC Pre-K funding exposure was larger for students who subsequently enrolled in school districts with lower rates of growth in academic achievement on average. Our findings provide support for dynamic substitutability as a mechanism by which educational environments during school might confer the long-term benefits of ECE. Such results emphasize that policies and decision-making around pre-K *cannot* be viewed in isolation from the educational inputs children receive in later grades. In the debate about long-term effects of ECE, we encourage educational stakeholders and policymakers to consider the role that subsequent educational environments can play in determining whether ECE effects are maintained or diminished in the long run. Attending to the quality of both ECE and later educational environments simultaneously is likely to lead to better outcomes for all students.

Endnotes

1. We refer to ECE program effects on eighth grade outcomes as “long-term” because students have entered adolescence. Effects on student outcomes during elementary school are often referred to as medium-term (e.g., Unterman & Weiland, 2020).
2. Between 2001 and 2010, the NC Pre-K program was administered by the North Carolina Department of Public Instruction (DPI) and known as the More at Four program. In 2011, the NC General Assembly transferred the program to the Division of Child Development and Early Education (DCDEE) in the North Carolina Department of Health and Human Services (DHHS) and the program was renamed the NC Pre-K program.
3. After our study period (i.e., through 2010), NC Pre-K program eligibility was extended to children with a parent serving in the military.
4. Subsequent years saw an increase in percentage of the children enrolled in a participating NC Pre-K classroom who received a funded slot increased, reaching 86% during the 2017-18 program year.
5. A school year is defined as July 1st of the previous year through June 30th of the listed year.
6. A total of 91 NC Pre-K program contractors served North Carolina’s 100 counties during each fiscal year. Funds awarded to a single contractor that served multiple counties were disaggregated across counties within the contractor’s service region based on the population of age-eligible children in each county.
7. School-district economic disadvantage was measured based on the percent of students in the school district who qualified for free- or reduced-price lunch.
8. The North Carolina Partnership for Children provided Smart Start funding data expended by program contractors in each of North Carolina’s 100 counties during each fiscal year.

9. The Smart Start and NC Pre-K funding exposure variables were based on the entire population of age-eligible children even if only a portion of children received direct financial support to participate in the program.
10. We note that 4 public school districts were closed and consolidated with other public-school districts during the course of our panel years. For our study, we aggregated information within these consolidated school districts prior to the years in which they were consolidated to utilize 115 school districts across all study years.
11. All charter school districts were excluded from our study. The following public-school districts were also excluded from our study because they were not comparable to the 115 traditional public-school districts: (1) the NC Schools for the Blind and Deaf and (2) the NC School of Science and Mathematics.
12. There are 75 counties in NC that were served by a single school district, while 15 counties were served by two school districts.
13. The following categories of students were excluded from our analyses because they took alternative assessments with test scales that were not equated with the EOG scale: (a) students with disabilities who took alternative assessments specifically designed to meet their needs and (b) advanced eighth grade students who were not assessed with the EOG Math test in 2019, but were instead assessed using the high-school end-of-course (EOC) math test (1.6% of our input sample).
14. The scores of individual students who were retained in grade and had valid test scores recorded for the same grade assessment in separate years were included in the models for multiple cohorts (e.g., the scores for a student who took the third-grade assessment during the 1997 and 1998 school years were included in the models for cohorts 2 and 3, respectively).

15. Within cohorts, the scores of students who changed school districts were included in each respective school district (e.g., the scores for a student who took the third- through fifth-grade assessments in school district A were nested within school district A, and the scores for that same student who took the six- through eighth-grade assessments in school district B were nested within school district B).
16. For students who repeated eighth grade, we used the test score that was recorded during the first school year in which they were enrolled in eighth grade.
17. In previous studies, we have found the pattern of findings remained the same between analyses based on county of residence at birth and county of school attendance (Bai et al., 2020).
18. This approach was employed in order to avoid the confound using student test score information from the same students as both independent and dependent variables, because regressing student's eighth grade achievement scores on school-district average achievement and achievement growth values measured from the same student test scores could introduce the same estimation errors on both the left- and right-hand side of the regression equation—resulting in biased estimates for school-district average achievement and achievement growth.
19. Please note that students in charter school districts were included in the transformation of EOG scale scores to Z-scores, but these students were excluded from Phase I and II analyses. Therefore, a Z-score value of 0.00 indicates the statewide average for students in public school districts and charter school districts. The negative mean values for school-district average achievement and achievement growth imply that students in charter school districts scored higher than students in public school districts.

20. The original values for school-district achievement growth represented a single grade/year of growth. These yearly values were rescaled to represent the linear rate of growth from third to eighth grade by multiplying the grade/year rate of growth produced in the Phase I analyses by a value of 5. The interpretation for a school-district achievement growth score of 0.00 is that the school district's rank-order standing remained at the statewide average from third to eighth grade. The interpretation for a positive school-district achievement growth score is that the school district gained in rank-order standing among all public-school districts from third to eighth grade, while the interpretation for a negative school-district achievement growth score is that the school district declined in rank-order standing among all public-school districts from third to eighth grade.
21. We conducted a series of endogeneity analyses to examine the association between NC Pre-K program funding exposure and school-district average achievement & achievement growth (see Table S4). We found no reliable association between NC Pre-K funding and school district average achievement or achievement growth in reading and math. These findings suggest that our analyses satisfy a necessary condition to examine moderation.
22. The 2009-10 program year was the most recent program year considered in our study, and this figure has remained relatively stable since then.

Table 1

Descriptive Statistics for the Eighth Grade Student Reading Achievement Analysis Sample (N = 1,160,725)

Variable	*M*	(*SD*)
NC Pre-K program funding		
NC Pre-K funding exposure (in 2019 $s; all study years)	420	(671)
Eighth grade achievement outcomes		
Reading	-0.01	(1.00)
Math	0.00	(1.00)
School-district average achievement		
Reading	-0.09	(0.18)
Math	-0.10	(0.19)
School-district achievement growth		
Reading	-0.02	(0.10)
Math	0.00	(0.16)
Student covariates		
Sex (Male; %)	50	
Economic disadvantage (%)	46	
Extremely low birth weight (%)	<1	
Very low birth weight (%)	1	
Low birth weight (%)	7	
Normal birth weight (%)	82	
High birth weight (%)	10	
Non-Hispanic White (%)	58	
African American (%)	30	
Hispanic (%)	7	
Native American (%)	1	
Asian (%)	2	
Mixed race (%)	3	
Quarter of birth 1 (%)	25	
Quarter of birth 2 (%)	25	
Quarter of birth 3 (%)	30	
Quarter of birth 4 (%)	21	
Maternal education (years)	12.58	(2.55)
Parent marital status (%)	65	
Maternal age (years)	26.11	(5.93)
No dad information on birth record (%)	14	
Mother immigrant (%)	9	
First born (%)	43	

Table 1 (continued)

Descriptive Statistics for the Eighth Grade Student Reading Achievement Analysis Sample (N = 1,160,725)

Variable	*M*	(*SD*)
Mother Non-Hispanic White (%)	60	
Mother African American (%)	29	
Mother Hispanic (%)	7	
Mother Native American (%)	2	
Mother Asian (%)	1	
Mother mixed race (%)	<1	
School-district covariates		
Total per pupil expenditures (in 2019 $s)	9822	(977)
Economic disadvantage (% in school district)	52.05	(15.76)
School district membership	39685	(45559)
County covariates		
Births to African American mothers (% of births)	25.74	(14.47)
Births to Hispanic mothers (% of births)	7.97	(11.07)
Births to low education mothers (% of births)	22.59	(6.56)
Population on Food Stamps (% of population)	7.41	(3.64)
Population on Medicaid (% of population)	14.42	(5.89)
Number of births (log transformed)	7.51	(1.09)
Total population (log transformed)	11.77	(1.00)
Median family income (in 2019 $s)	68309	(13232)
Smart Start funding exposure (in 2019 $s; all study years)	1462	(1088)

Note. Descriptive statistics are provided for those students included in the reading outcome analyses. Descriptive statistics for students included in the math outcome analyses are provided in Table S3. Dollar amounts for NC Pre-K funding exposure, total per pupil expenditures, median family income, and Smart Start funding exposure variables were scaled to raw dollar amounts in fiscal year 2019 $s, with all study years represented (including pre-trend years in which NC Pre-K funding exposure was equal to 0 for students in all counties). School-district average achievement in reading/math was measured at third grade and achievement growth in reading/math was measured from third to eighth grade. All values of school-district average achievement and achievement growth were “current year minus 1.”

Table 2

Regression Results: Linear and Quadratic Main Effects of NC Pre-K Funding Exposure

	Reading				Math			
	Linear Model		Quadratic Model		Linear Model		Quadratic Model	
Parameter	β	(*SE*)	β	(*SE*)	β	(*SE*)	β	(*SE*)
NC Pre-K funding exposure	0.08***	(0.02)	0.09*	(0.04)	0.11*	(0.04)	0.06	(0.08)
NC Pre-K funding exposure²			0.00	(0.02)			0.03	(0.03)
R-square	0.27		0.27		0.27		0.27	

Note. Sample sizes for the reading and math outcome analyses are 1,160,725 and 1,123,347, respectively. Parameter estimates for the covariates are not displayed in Table 2, but can be made available from the corresponding author upon request (including parameter estimates for the student, school-district, and county covariates; county fixed effects; program year fixed effects; and linear time trends). The effect of NC Pre-K funding exposure is equal to a change from $0 to $2,030. Eighth grade reading and math achievement were grand mean standardized with $M = 0$, $SD = 1$. All models were estimated with robust standard errors clustered at the county level. $^{\dagger}p < .10$, $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$

Table 3

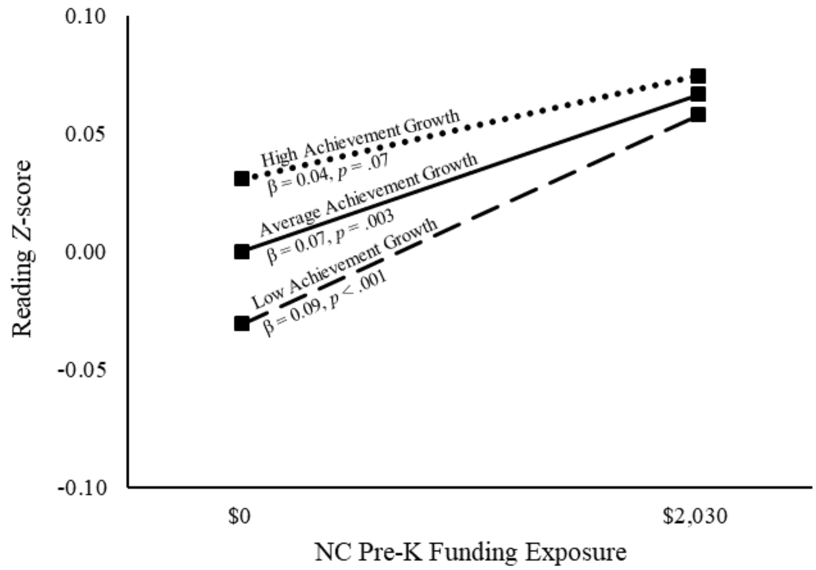
Eighth-Grade Student Achievement Analysis Results

	Reading		Math	
Parameter	β	(*SE*)	β	(*SE*)
NC Pre-K funding exposure	0.07**	(0.02)	0.09*	(0.04)
Average Achievement	0.04***	(0.01)	0.07***	(0.01)
Achievement Growth	0.03***	(0.00)	0.08***	(0.00)
NC Pre-K × Average Achievement	–0.01	(0.01)	–0.01	(0.01)
NC Pre-K × Achievement Growth	–0.02***	(0.01)	–0.03**	(0.01)
R-squared	0.27		0.27	

Note. Sample sizes for the reading and math outcome analyses are 1,160,725 and 1,123,347, respectively. Parameter estimates for the covariates are not displayed in Table 3, but can be made available from the corresponding author upon request (including parameter estimates for the student, school-district, and county covariates; county fixed effects; program year fixed effects; and linear time trends). School-district average achievement/achievement growth correspond to reading average achievement/achievement growth in the reading models and math average achievement/achievement growth in the math models. All values of school-district average achievement and achievement growth were "current year minus 1." The effect of NC Pre-K funding exposure is equal to a change from \$0 to \$2,030. Eighth grade reading and math achievement as well as school district average achievement and achievement growth were grand mean standardized with $M = 0$, $SD = 1$. All models were estimated with robust standard errors clustered at the county level. $^{\dagger}p < .10$, $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$

Figure 1

Moderation of the NC Pre-K Funding Effect on Eighth Grade Student Reading Achievement by School-District Achievement Growth in Reading



Note. The effect of NC Pre-K funding exposure on student reading achievement in eighth grade at three values of school-district achievement growth in reading: 1 *SD* below the average (low), at the average, and 1 *SD* above the average (high).

References

Abenavoli, R. M. (2019). The mechanisms and moderators of “fade-out”: Towards understanding why the skills of early childhood program participants converge over time with the skills of other children. *Psychological Bulletin*, *145*(12), 1103–1127. https://doi.org/10.1037/bul0000212

Andrabi, T., Das, J., Ijaz Khwaja, A., & Zajonc, T. (2011). Do value-added estimates add value? Accounting for learning dynamics. *American Economic Journal: Applied Economics*, *3*(3), 29–54. https://doi.org/10.1257/app.3.3.29

Bai, Y., Ladd, H. F., Muschkin, C. G., & Dodge, K. A. (2020). Long-term effects of early childhood programs through eighth grade: Do the effects fade out or grow? *Children and Youth Services Review*, *112*(1), 1–10. https://doi.org/10.1016/j.childyouth.2020.104890

Bailey, D. H., Duncan, G. J., Cunha, F., Foorman, B. R., & Yeager, D. S. (2020). Persistence and fade-out of educational-intervention effects: Mechanisms and potential solutions. *Psychological Science in the Public Interest*, *21*(2), 55–97. https://doi.org/10.1177/1529100620915848

Bailey, D. H., 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*(1), 7–39. https://doi.org/10.1080/19345747.2016.1232459

Bailey, D. H., Jenkins, J. M., & Alvarez-Vargas, D. (2020). Complementarities between early educational intervention and later educational quality? A systematic review of the sustaining environments hypothesis. *Developmental Review*, *56*(1), 1–22. https://doi.org/10.1016/j.dr.2020.100910

Bailey, M. J., Sun, S., & Timpe, B. (2021). Prep school for poor kids: The Long-run impacts of Head Start on human capital and economic self-sufficiency. *American Economic Review*, *111*(12), 3963–4001. https://doi.org/10.1257/aer.20181801

Barnett, W. S., Epstein, D. J., Carolan, M. E., Fitzgerald, J., Ackerman, D. J., & Friedman, A. H. (2010). *The State of Preschool 2010: State Preschool Yearbook*. National Institute for Early Education Research. https://nieer.org

Barnett, W. S., Hustedt, J. T., Hawkinson, L. E., & Robin, K. B. (2006). *The state of preschool 2006: State preschool yearbook*. National Institute for Early Education Research. https://nieer.org

Barnett, W. S., Robin, K. B., Hustedt, J. T., & Schulman, K. L. (2003). *The state of preschool 2003: State preschool yearbook*. National Institute for Early Education Research. https://nieer.org

Barr, A., & Gibbs, C. R. (2022). Breaking the Cycle? Intergenerational effects of an antipoverty program in early childhood. *Journal of Political Economy*, *130*(12), 3253–3285. https://doi.org/10.1086/720764

Bazemore, M., & Van Dyk, P. B. (2004). *North Carolina reading comprehension tests: Technical report (Ed.2)*. North Carolina Department of Public Instruction.

Blazar, D., & Schueler, B. (2022). Why do school districts matter? An interdisciplinary framework and empirical review. (EdWorkingPaper: 22-581). Retrieved from Annenberg Institute at Brown University: https://doi.org/10.26300/58m4-fs65

Bouguen, A., Filmer, D., Macours, K., & Naudeau, S. (2018). Preschool and parental response in a second best world: Evidence from a school construction experiment. *Journal of Human Resources*, *53*(2), 474–512. https://doi.org/10.3368/jhr.53.2.1215-7581R1

Bryk, A. S., & Raudenbush, S. W. (1992). *Hierarchical linear models: Applications and data analysis methods*. Sage Publications, Inc.

Bryk, A. S., Sebring, P. B., Allensworth, E., Luppescu, S., & Easton, J. Q. (2010). *Organizing schools for improvement: Lessons from Chicago*. The University of Chicago Press.

Burchinal, M., Foster, T., Garber, K., Cohen-Vogel, L., Bratsch-Hines, M., & Peisner-Feinberg, E. (2022). Examining three hypotheses for pre-kindergarten fade-out. *Developmental Psychology, 58*(3), 453–469. https://doi.org/10.1037/dev0001302

Burchinal, M., Whitaker, A., Jenkins, J., Bailey, D., Watts, T., Duncan, G., & Hart, E. (2024). Unsettled science on longer-run effects of early education. *Science*, *384*(6695), 506–508. https://doi.org/10.1126/science.adn2141

Callaway, B., Goodman-Bacon, A., & Sant'Anna, P. H. (2024). Difference-in-differences with a continuous treatment. *NBER Working Paper Series*, *Working Paper 32117*. https://doi.org/10.48550/arXiv.2107.02637

Carr, R. C. (2021). *The benefits of early childhood education can persist in the long run*. The Hunt Institute.

Carr, R. C., Jenkins, J. M., Watts, T. W., Peisner-Feinberg, E. S., & Dodge, K. A. (2024). Investigating if high-quality kindergarten teachers sustain the pre-K boost to children's emergent literacy skill development in North Carolina. *Child Development*, *95*(4), 1200–1217. https://doi.org/10.1111/cdev.14076

Carr, R. C., & Peisner-Feinberg, E. S. (2021, September 13). High-quality elementary schools solidify the benefits of NC Pre-K. *EducationNC*. Retreived from www.ednc.org

Carr, R. C., Peisner-Feinberg, E. S., Kaplan, R., & Mokrova, I. L. (2021). Effects of North Carolina's pre-kindergarten program at the end of kindergarten: Contributions of school-wide quality. *Applied Developmental Psychology*, *76*(1), 1–12. https://doi.org/10.1016/j.appdev.2021.101317

Cascio, E. U. (2021). Early childhood education in the United States. In *The Routledge Handbook of the Economics of Education* (pp. 30–72). https://doi.org/10.4324/9780429202520-3

Cascio, E. U., & Schanzenbach, D. W. (2013). The impacts of expanding access to high-quality preschool education. *Brookings Papers on Economic Activity*, *47*(2 (Fall), 127–192. https://www.brookings.edu/bpea-articles/

Chenoweth, K. (2017). *Schools that succeed: How educators marshal the power of systems for improvement*. Harvard Education Press.

Chenoweth, K. (2021). *Districts that succeed: Breaking the correlation between race, poverty, and achievement*. Harvard Education Press.

Cohen-Vogel, L., Sadler, J., Little, M. H., Merrill, B., & Curran, F. C. (2020). The adoption of public pre-kindergarten among the American states: An event history analysis. *Educational Policy*, *36*(6), 1227–1577. https://doi.org/10.1177/0895904820961002

Cunha, F., Heckman, J., Lochner, L., & Masterov, D. V. (2006). Interpreting the evidence of life cycle skill formation. In F. Hanushek & F. Welch (Eds.), *Handbook of the economics of education* (pp. 698–812). Elsevier.

Dodge, K. A., Bai, Y., Ladd, H. F., & Muschkin, C. G. (2016). Impact of North Carolina's early childhood programs and policies on educational outcomes in elementary school. *Child Development*, *88*(3), 996–1014. https://doi.org/10.1111/cdev.12645

Dynarski, S., Jacob, B., & Kreisman, D. (2018). How important are fixed effects and time trends in estimating returns to schooling? Evidence from a replication of Jacobson, Lalonde, and

Sullivan, 2005. *Journal of Applied Econometrics*, *33*(7), 1098–1108. https://doi.org/10.1002/jae.2653

Eccles, J. S., & Roeser, R. W. (2010). An ecological view of schools and development. In J. S. Eccles & J. L. Meece (Eds.), *Handbook of research on schools, schooling, and human development* (pp. 6–21). Routledge.

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

Friedman-Krauss, A. H., Barnett, S. W., Garver, K. A., Hodges, K. S., Weisenfeld, G. G., & Gardiner, B. A. (2020). *The state of preschool 2019: State preschool yearbook*. National Institute for Early Education Research. https://nieer.org

Friedman-Krauss, A. H., Barnett, S. W., Hodges, K. S., Garver, K. A., Jost, T. M., Weisenfeld, G. G., & J., D. (2024). *The state of preschool 2023: State preschool yearbook*. National Institute for Early Education Research. https://nieer.org

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, *225*(2), 254–277. https://doi.org/10.1016/j.jeconom.2021.03.014

Heckman, J. J. (2006). Skill Formation and the Economics of Investing in Disadvantaged Children. *Science*, *312*(5782), 1900–1902. https://doi.org/10.1126/science.1128898

Johnson, R. C., & Jackson, C. K. (2019). Reducing inequality through dynamic complementarity: Evidence from Head Start and public school spending. *American Economic Journal: Economic Policy*, *11*(4), 310–349. https://doi.org/10.1257/pol.20180510

Kauerz, K., & Coffman, J. (2013). *Framework for planning, implementing, and evaluating prek-3rd grade approaches*. National P-3 Center.

Kose, E. (2021). Public investments in early childhood education and academic performance: Evidence from Head Start in Texas. *Journal of Human Resources*, *58*(6). https://doi.org/10.3368/jhr.0419-10147R2

Kraft, M. A. (2020). Interpreting effect sizes of education interventions. *Educational Researcher*, *49*(4), 241–253. https://doi.org/10.3102/0013189x20912798

Kropko, J., & Kubinec, R. (2020). Interpretation and identification of within-unit and cross-sectional variation in panel data models. *PLoS ONE*, *15*(4), 1–22. https://doi.org/10.1371/journal.pone.0231349

Ladd, H. F., Muschkin, C. G., & Dodge, K. A. (2013). From birth to school: Early childhood initiatives and third-grade outcomes in North Carolina. *Journal of Policy Analysis and Management*, *33*(1), 162–187. https://doi.org/10.1002/pam.21734

Lipsey, M. W., Farran, D. C., & Durkin, K. (2018). Effects of the Tennessee Prekindergarten Program on children's achievement and behavior through third grade. *Early Childhood Research Quarterly*, *45*(1), 155–176. https://doi.org/10.1016/j.ecresq.2018.03.005

Martinez, S., Naudeau, S., & Pereira, V. A. (2017). Preschool and child development under extreme poverty: Evidence from a randomized experiment in rural Mozambique. *World Bank Policy Research Working Paper*, *0*(8290), 1–48. https://doi.org/10.1596/1813-9450-8290

Mbella, K., Zhu, M., Karkee, T., & Lung, H. (2016a). *The North Carolina testing program technical report, 2012–2015: English language arts/reading assessments (ELA) end-of-grade 3–8 and end-of-course English II*. Department of Public Instruction.

Mbella, K., Zhu, M., Karkee, T., & Lung, H. (2016b). *The North Carolina testing program technical report, 2012–2015: Mathematics assessments end-of-grade 3–8 and end-of-course Math I*. Department of Public Instruction.

Muschkin, C. G., Ladd, H. F., & Dodge, K. A. (2015). Impact of North Carolina's early childhood initiatives on special education placements in third grade. *Educational Evaluation and Policy Analysis*, *37*(4), 478–500. https://doi.org/10.3102/0162373714559096

Muschkin, C. G., Ladd, H. F., & Sauval, M. (2024). Pre-K enrollments and teaching environments in North Carolina elementary schools. *Children and Youth Services Review*, *164*(1), 1–11. https://doi.org/https://doi.org/10.1016/j.childyouth.2024.107832

North Carolina Department of Public Instruction. (2009). *North Carolina reading comprehension tests: Technical report (Ed. 3)*. Author.

North Carolina Partnership for Children. (2019). *Annual report to the North Carolina General Assembly for fiscal year 2018-2019*. https://www.smartstart.org

Peisner-Feinberg, E., Kuhn, L., Zadrozny, S., Foster, T., & Burchinal, M. (2020). *Kindergarten follow-up findings from a small-scale RCT study of the North Carolina Pre-Kindergarten Program*. The University of North Carolina at Chapel Hill.

Peisner-Feinberg, E. S., Elander, K. C., Maris, C. L., & the More at Four Evaluation Team. (2006). *Evaluation of the North Carolina More at Four Pre-Kindergarten Program, year 4.* The University of North Carolina, Frank Porter Graham Child Development Institute.

Peisner-Feinberg, E. S., Schaaf, J., Hildebrandt, L., & LaForett, D. (2013). *Quality and characteristics of the North Carolina Pre-Kindergarten program: 2011-2012 statewide evaluation*. The University of North Carolina, FPG Child Development Institute.

Peisner-Feinberg, E. S., & Schaaf, J. M. (2010). *Long-term effects of the North Carolina More at Four Pre-kindergarten Program: Children's reading and math skills at third grade: Full report*. The University of North Carolina, FPG Child Development Institute.

Peisner-Feinberg, E. S., & Schaaf, J. M. (2011). *Effects of the North Carolina More at Four Pre-kindergarten Program on children's school readiness skills: Summary of key findings*. The University of North Carolina, FPG Child Development Institute.

Peisner-Feinberg, E. S., Zadrozny, S., Kuhn, L., & Van Manen, K. (2019). *Effects of the North Carolina Pre-Kindergarten Program: Findings through Pre-K of a Small-Scale RCT Study*. The University of North Carolina, FPG Child Development Institute.

Phillips, D. A., Lipsey, M. W., Dodge, K. A., Haskins, R., Bassok, D., Burchinal, M. R., Duncan, G. J., Dynarski, M., Magnuson, K. A., & Weiland, C. (2017). *The current state of scientific knowledge on pre-kindergarten effects.* Brookings Institution.

Puma, M., Bell, S., Cook, R., Heid, C., Shapiro, G., Broene, P., Jenkins, F., Fletcher, P., Quinn, L., Friedman, J., Ciarico, J., Rohacek, M., Adams, G., & Spier, E. (2010). *Head Start Impact Study: Final report*. Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.

Reardon, S. F. (2019). Educational opportunity in early and middle childhood: Using full population administrative data to study variation by place and age. *RSF: The Russell Sage Foundation Journal of the Social Sciences, 5*(2), 40–68. https://doi.org/10.7758/rsf.2019.5.2.03

Sanford, E. E. (1996). *North Carolina end-of-grade tests technical report # 1: Reading comprehension, mathematics*. Department of Public Instruction.

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. Oxford University Press.

Thompson, O. (2018). Head Start's long-run impact. *Journal of Human Resources*, *53*(4), 1100–1139. https://doi.org/10.3368/jhr.53.4.0216.7735R1

Watts, T. W., Jenkins, J. M., Dodge, K. A., Carr, R. C., Sauval, M., Bai, Y., Escueta, M., Duer, J., Ladd, H., Muschkin, C., Peisner-Feinberg, E., & Ananat, E. (2023). Understanding heterogeneity in the impact of public preschool programs. *Monographs of the Society for Research in Child Development*, *88*(1), 1–182. https://doi.org/10.1111/mono.12463

Williams, B. J. (2019). The spillover benefits of expanding access to preschool. *Economics of Education Review*, *70*, 127–143. https://doi.org/10.1016/j.econedurev.2019.04.002

Yoshikawa, H., Weiland, C., Brooks-Gunn, J., Burchinal, M. R., Espinosa, L. M., Gromley, W. T., Ludwig, J., Magnuson, K. A., Phillips, D., & Zaslow, M. J. (2013). *Investing in our future: The evidence base on preschool education*. Foundation for Child Development.

Online Supplementary Files

Table S1

Descriptive Statistics for School-District Average Achievement and Achievement Growth Presented at the School-District Level (N = 2,070)

Variable	*M*	(*SD*)
School-district average achievement		
Reading	–0.13	(0.22)
Math	–0.14	(0.23)
School-district achievement growth		
Reading	–0.02	(0.12)
Math	–0.01	(0.20)

Note. School-district average achievement in reading/math (β_{00j}) was measured at third grade and achievement growth in reading/math (β_{10j}) was measured from third to eighth grade. Economic disadvantage was measured based on the percent of students in the school district who qualified for free- or reduced-price lunch. There were 115 school districts in each of the 18 cohorts (i.e., 18 × 15 = 2,070 school district by cohort observations).

Table S2

Correlates of School-District Average Achievement & Achievement Growth (N = 2,070)

	1	2	3	4	5
School-district average achievement					
1. Reading	1.00				
2. Math	0.92^{***}	1.00			
School-district achievement growth					
3. Reading	0.04^{*}	$0.04^{\dagger}$	1.00		
4. Math	0.03	$–0.16^{***}$	0.66^{***}	1.00	
Economic Disadvantage					
5. Economic disadvantage (%)	$–0.65^{***}$	$–0.65^{***}$	$–0.19^{***}$	$–0.07^{**}$	1.00

Note. Economic disadvantage is the percent of economically disadvantaged students in the school district (i.e., the percent of students in the school district who qualified for free- or reduced-price lunch). $^{\dagger}p < .10$, $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$

Table S3

Descriptive Statistics for the Eighth Grade Student Math Achievement Analysis Sample (N = 1,123,347)

Variable	*M*	(*SD*)
NC Pre-K program funding		
NC Pre-K funding exposure (in 2019 $s; all study years)	385	(640)
Eighth grade achievement outcomes		
Reading	-0.04	(1.00)
Math	0.00	(1.00)
School-district average achievement		
Reading	-0.09	(0.18)
Math	-0.10	(0.19)
School-district achievement growth		
Reading	-0.02	(0.10)
Math	-0.01	(0.16)
Student covariates		
Sex (Male; %)	50	
Economic disadvantage (%)	47	
Extremely low birth weight (%)	<1	
Very low birth weight (%)	1	
Low birth weight (%)	7	
Normal birth weight (%)	82	
High birth weight (%)	10	
Non-Hispanic White (%)	58	
African American (%)	30	
Hispanic (%)	7	
Native American (%)	1	
Asian (%)	2	
Mixed race (%)	3	
Quarter of birth 1 (%)	24	
Quarter of birth 2 (%)	25	
Quarter of birth 3 (%)	30	
Quarter of birth 4 (%)	21	
Maternal education (years)	12.53	(2.53)
Parent marital status (%)	65	
Maternal age (years)	26.01	(5.91)
No dad information on birth record (%)	14	
Mother immigrant (%)	9	
First born (%)	43	

Table S3 (continued)

Descriptive Statistics for the Eighth Grade Student Math Achievement Analysis Sample (N = 1,123,347)

Variable	*M*	(*SD*)
Mother Non-Hispanic White (%)	61	
Mother African American (%)	29	
Mother Hispanic (%)	7	
Mother Native American (%)	2	
Mother Asian (%)	1	
Mother mixed race (%)	<1	
School-district covariates		
Total per pupil expenditures (in 2019 $s)	9817	(979)
Economic disadvantage (% in school district)	51.93	(15.64)
School district membership	38871	(44675)
County covariates		
Births to African American mothers (% of births)	25.84	(14.54)
Births to Hispanic mothers (% of births)	7.71	(11.12)
Births to low education mothers (% of births)	22.60	(6.58)
Population on Food Stamps (% of population)	7.39	(3.65)
Population on Medicaid (% of population)	14.35	(5.89)
Number of births (log transformed)	7.49	(1.08)
Total population (log transformed)	11.75	(1.00)
Median family income (in 2019 $s)	68018	(13141)
Smart Start funding exposure (in 2019 $s; all study years)	1452	(1103)

Note. Descriptive statistics are provided for those students included in the math outcome analyses. Dollar amounts for NC Pre-K funding exposure, total per pupil expenditures, median family income, and Smart Start funding exposure variables were scaled to raw dollar amounts in fiscal year 2019 $s, with all study years represented (including pre-trend years in which NC Pre-K funding exposure was equal to 0 for students in all counties). School-district average achievement in reading/math was measured at third grade and achievement growth in reading/math was measured from third to eighth grade. All values of school-district average achievement and achievement growth were “current year minus 1.”

Table S4

Analyses to Examine the Association between Program Funding Exposure and School-District Average Achievement & Achievement Growth (N = 2,070)

	Reading				Math			
	Average Achievement		Achievement Growth		Average Achievement		Achievement Growth	
Parameter	*B*	(*SD*)	*B*	(*SD*)	*B*	(*SD*)	*B*	(*SD*)
NC Pre-K funding exposure	0.00	(0.02)	0.03	(0.02)	0.00	(0.03)	–0.01	(0.05)

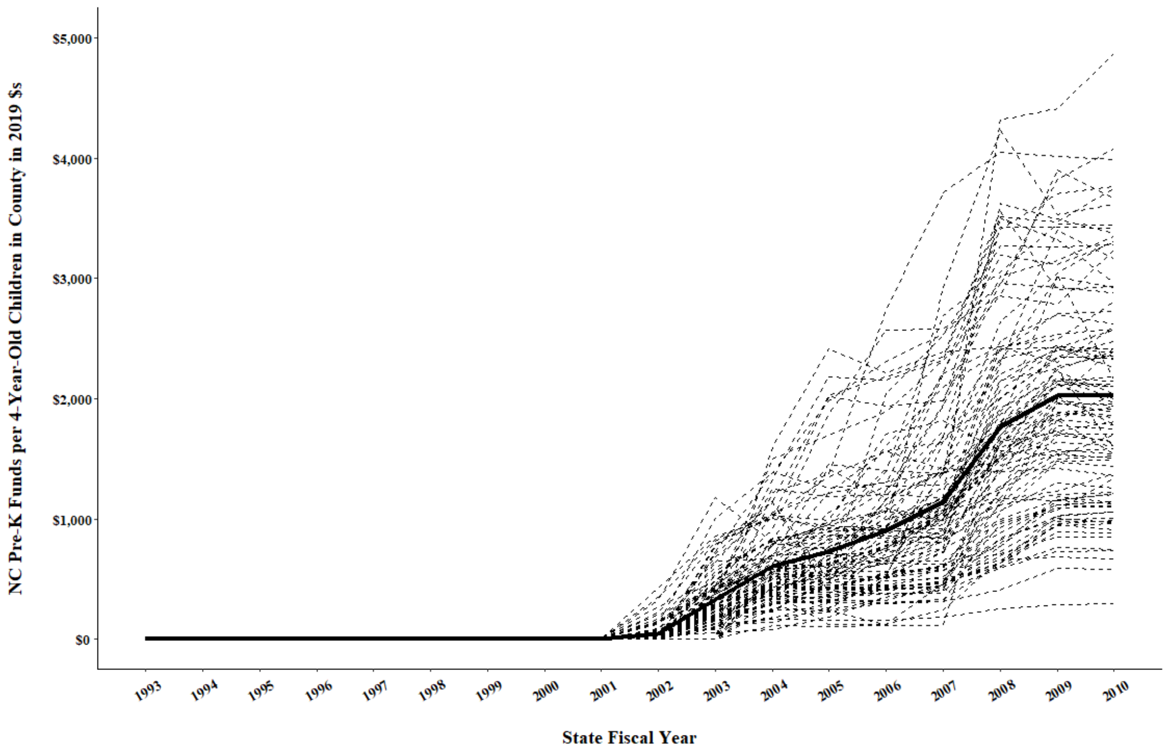
Note. The generalized equation for these analyses was:

$$O_{sPy+4} = \beta_0 + \beta_1 NC\ PreK\ funding_{cPy} + \alpha_c + \gamma_{Py} + \alpha * \gamma_{cPy} + \epsilon_{scPy}$$

O is the school-district average achievement or achievement growth score for school district *s* in county *c* aligned to program year *Py* + 4, β_1 is the main effect of NC Pre-K funding in county *c* and program year *Py*, α is the fixed effect for county, and γ is the fixed effect for program year. School-district achievement and school-district average achievement growth information was aligned to county-level program funding information in program year *Py* + 4 (e.g., students' exposure to NC Pre-K program funding during program year 2000 was aligned to the average achievement and achievement growth values for the cohort of students enrolled in third grade four years later during school year 2004, which is the school year in which these students were expected to be in third grade), and all values of school-district average achievement and achievement growth were "current year minus 1", which was consistent with the approach used in our eighth-grade student achievement analyses. The effect of NC Pre-K funding exposure is equal to a change from \$0 to \$2,030. All models were estimated with robust standard errors clustered at the county level. $^{\dagger}p < .10$, $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$

Figure S1

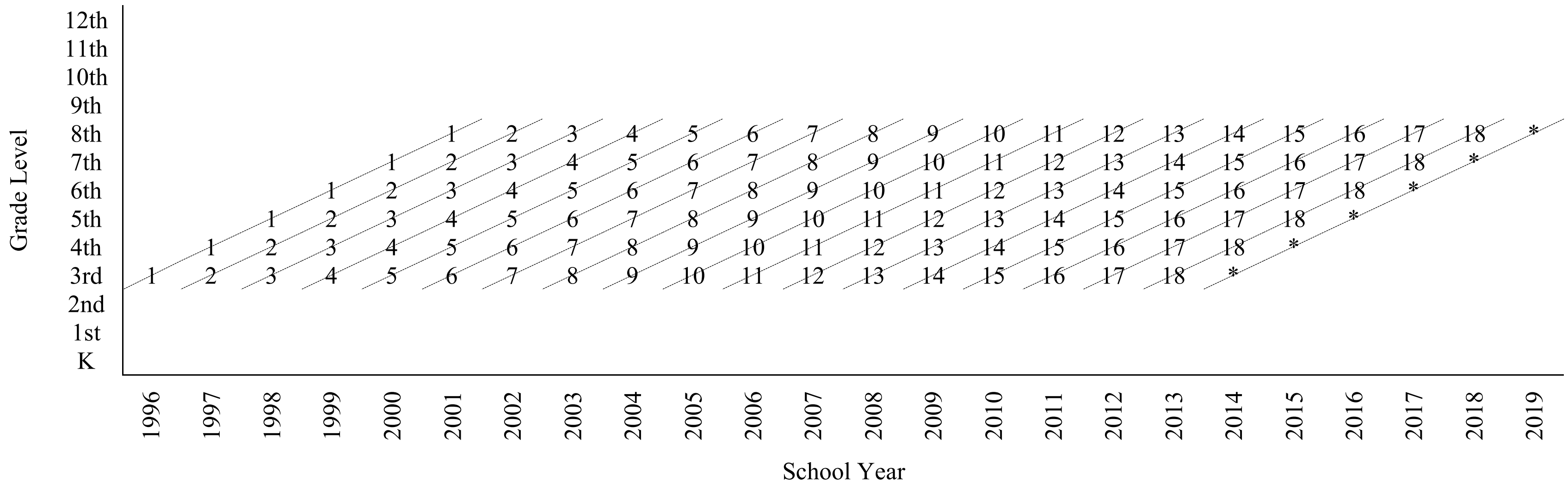
NC Pre-K Funds per 4-Year-Old Children in County: 1993 to 2010



Note. All counties with statewide average. Dotted lines represent individual counties. Solid line represents statewide average. NC Pre-K was established in state fiscal year 2002. In the student outcome analyses, all counties were assigned a pre-trend value of "0" prior to the establishment of NC Pre-K in state fiscal year 2002.

Figure S3

North Carolina Third to Eighth Grade Student Cohorts between 1996 and 2018 (N = 18)



Note. Each diagonal line corresponds to a cohort of students. In Phase I of our analyses, a separate growth curve model was estimated for students in each cohort, which was used to calculate the intercept (i.e., third grade average achievement) and slope (i.e., achievement growth between third and eighth grade) in reading and mathematics for each school district within each cohort. These estimates were then used in Phase II of our analyses to examine moderation of the long-term effect of NC Pre-K. Students in the cohort indexed by the asterisk were included in the Phase II analyses, but not the Phase I analyses because of the "current year minus 1" approach.