Monographs of the Society for Research in Child Development
Ginger A. Moore, Editor

Understanding Heterogeneity in the Impact of Public Preschool Programs

Tyler W. Watts,[1] Jade M. Jenkins,[2] Kenneth A. Dodge,[3] Robert C. Carr,[4] Maria Sauval,[2] Yu Bai,[3] Maya Escueta,[3] Jennifer Duer,[5] Helen Ladd,[3] Clara Muschkin,[3] Ellen Peisner-Feinberg,[4] and Elizabeth Ananat[6]

Abstract We examine the North Carolina Pre-K (NC Pre-K) program to test the hypothesis that observed variation in effects resulting from exposure to the program can be attributed to interactions with other environmental factors that occur before, during, or after the pre-k year. We examine student outcomes in 5th grade and test interaction effects between NC's level of investment in public pre-k and moderating factors. Our main sample includes the population of children born in North Carolina between 1987 and 2005 who later attended a public school in that state, had valid achievement data in 5th grade, and could be matched by administrative record review ($n = 1{,}207{,}576$; 58% White non-Hispanic, 29% Black non-Hispanic, 7% Hispanic, 6% multiracial and Other race/ethnicity). Analyses were based on a natural experiment leveraging variation in county-level funding for NC Pre-K across NC counties during each of the years the state scaled up the program. Exposure to NC Pre-K funding was defined as the per-4-year-old-child state allocation of funds to a county in a year. Regression models included child-level and county-level covariates and county and year fixed effects.

Estimates indicate that a child's exposure to higher NC Pre-K funding was positively associated with that child's academic achievement 6 years later. We found no effect on special education placement or grade retention. NC Pre-K funding effects on achievement were positive for all subgroups tested, and statistically significant for most. However, they were larger for children exposed to more disadvantaged environments either before or after the pre-k experience, consistent with a compensatory model where pre-k provides a

[1]Teachers College, Columbia University, [2]University of California, Irvine, [3]Duke University, [4]University of North Carolina at Chapel Hill, [5]Rutgers University, [6]Barnard College, Columbia University

Corresponding author: Kenneth A. Dodge, Duke University, email: dodge@duke.edu

Citation Information: 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*, *87*(4).

https://doi.org/10.1111/mono.12463

buffer against the adverse effects of prior negative environmental experiences and protection against the effects of future adverse experiences. In addition, the effect of NC Pre-K funding on achievement remained positive across most environments, supporting an additive effects model. In contrast, few findings supported a dynamic complementarity model. Instrumental variables analyses incorporating a child's NC Pre-K enrollment status indicate that program attendance increased average 5th grade achievement by approximately 20% of a standard deviation, and impacts were largest for children who were Hispanic or whose mothers had less than a high school education. Implications for the future of pre-k scale-up and developmental theory are discussed.

I. Study Introduction and Overview

When universal public education was first contemplated in the United States 200 years ago, the age at which children were required to matriculate was set at 7 or 8 years. This decision was not arbitrary, as this was considered the “age of reason,” meaning that it was understood by leading educators and developmental scientists of the day to be the age at which children first begin to think critically (Sameroff & Haith, 1996). Education prior to that age was assumed to be pointless because it was thought that teaching younger children was impossible, learning would be forgotten, or that noneducated children would naturally catch up when they arrived at elementary school. Since that time, the matriculation age has steadily moved downward. Kindergarten (literally, “children's garden”) was introduced 150 years ago as a play-oriented experience following a German model predicated on the belief that 5-year-old children cannot reason or think critically, but could blossom through play (Wollons, 2000). Education-oriented kindergarten began only more recently, and while it is now compulsory in 19 states, most of the nation offers it voluntarily and to great demand (Education Commission of the States, 2014). Today, the majority of 5-year-olds in the United States attend part-day or full-day kindergarten, and about one third of 4-year-olds attend some kind of publicly-funded educational program (Cascio, 2021).

Though the long-term effects of this uptick in enrollment in early educational programs remain uncertain, developmental scientists have long argued that early cognitive stimulation, whether in the home or through formal educational experiences, may be crucial to positive lifelong outcomes. We now know that young children's brains develop at a phenomenal rate during the first several years of life (Purves & Lichtman, 1985). The first 5 years of life are a sensitive period in development when young children are particularly receptive to learning new associations that endure across the lifespan and guide future learning (Brown & Jernigan, 2012). Theories of skill-building have described sequences of learning wherein the acquisition of initial skills are required to lay the foundation for more complicated skill learning (Taie, 2014). A child who misses out on initial skill acquisition may never catch up if subsequent teaching assumes prerequisite skills and focuses exclusively on more complicated skills.

Indeed, empirical evidence suggests that skill building may be especially dependent on environmental inputs during the earliest years (Huttenlocher et al., 1998; Magnuson & Waldfogel, 2008), and these inputs also shape disparities across groups in kindergarten readiness. Early socioeconomic status (SES) contributes to large academic skill disparities; trends that are present by the time children matriculate in kindergarten and never disappear (Garcia & Weiss, 2017). Children at the 10th percentile in SES begin kindergarten far behind and remain 3–4 years behind those at the 90th percentile, presumably because low SES deprives children of opportunities for cognitive stimulation and access to educational resources (Hanushek et al., 2019).

Researchers have turned to early childhood educational programs as a potential means of intervention to help address these early disparities. Rigorous randomized controlled trials (RCTs) of model early childhood programs with small numbers of children from low-income backgrounds have shown that early educational interventions to support parents at home and teach children in educational settings can have a long-term positive impact on important outcomes lasting into adulthood (Campbell et al., 2014; Schweinhart et al., 2005). Drawing heavily from these findings, developmental science has made a strong case for education-oriented early intervention, especially pre-kindergarten (pre-k)[1] (Yoshikawa et al., 2013). Public policymakers in the United States have heard this message. In 1989, the National Governors Association established a series of national education goals, the first of which focused on "the readiness of children to start school" (Vinovskis, 1999). President George H. W. Bush later reaffirmed this goal in his 1990 State of the Union address: "By the year 2000, every child must start school ready to learn" (Bush, 1990). Indeed, the share of 4-year-old children enrolled in public pre-k programs saw a dramatic increase, from 10% to 37% between 1978 and 2015 (Cascio, 2021), with much of this increase being driven by increased enrollment in state-funded pre-k programs (e.g., Barnett et al., 2017).

Debate continues about the expansion of public pre-k education. President Joseph Biden proposed large federal allocations for states to expand pre-k and implement other early childhood initiatives (Blad, 2021). Some critics have pushed back by questioning whether the government should support early education and, more pertinently, whether the scientific evidence indicates that pre-k has positive, null, or negative long-term outcomes (Lucey & Parti, 2021).

In light of that controversy, the early childhood education (ECE) field has been relatively rich with data on impact for children. Beginning with initial legacy studies (reviewed below), through numerous evaluations of the federal Head Start (HS) program, and carrying into today's evaluations, there is a consensus that empirical evaluations of the impact of scaled-up pre-k show robust average positive effects on children's academic skills immediately at the conclusion of the pre-k year (Phillips & Pre-Kindergarten Task Force, 2017). Children randomly assigned or selectively exposed to pre-k begin kindergarten at higher skill levels, on average, when compared with similar children who did not attend pre-k.

However, the field has yet to reach a consensus regarding how pre-k effects unfold after program participation ends. Some recent work has found that children exposed to higher levels of funding for pre-k continue to display higher achievement than children exposed to lower levels of funding through 8th grade (Bai et al., 2020). Conversely, other studies show that positive impacts from pre-k assignment fade out, or even become adverse, during elementary school (Lipsey et al., 2018). We define fadeout as the disappearing difference over time in early intervention impact between the "intervention" and "control" group in an ECE evaluation study. Recent work on fadeout suggests that much of the impact loss in the early intervention group is better characterized as "catch-up" by children in the control group in subsequent learning environments (Kang et al., 2019).

One robust conclusion from the body of evidence is that the measured long-term effects of experiencing pre-k are heterogeneous. Indeed, Raudenbush and Bloom (2015) argue that much more educational research should focus on the study of impact heterogeneity. Whitehurst (2018) has called for rigorous study of heterogeneous impacts to guide future policy, redesign interventions, or re-target them to different children. Indeed, the field has placed increasing interest on the ways that impact estimates vary, with sources of heterogeneity involving both characteristics of the child and their environments. For example, some work has examined how impacts differ based on demographic characteristics, such as the education level of children's mothers (Walters, 2015) and whether English is spoken in a child's home (Bloom & Weiland, 2015). Others have suggested that characteristics of the pre-k program itself or co-occurring environmental experiences during the preschool years may hold the key for understanding whether pre-k effects will last in the long term. Perhaps most relevant for the study of longer-term effects, others point to heterogeneity in environmental experiences that children encounter *after* the pre-k year ends. Bailey et al.'s (2017) seminal paper offers a framework for understanding fadeout; the authors call for closer scrutiny of a child's subsequent environments, including the child's school, teacher, and peer environments, as well as the community economic standing and child-specific home characteristics.

Study Overview

The goal of this Monograph is to understand heterogeneity in the long-term impacts of a scaled-up statewide pre-k program, North Carolina's Pre-K (NC Pre-K) program that has demonstrated overall positive main effects. By identifying factors associated with long-term positive impact, we hope to inform theories of early learning, help pre-k program administrators improve their services, and guide policymakers on how to allocate future funding.

We assert that although some differences in the impact estimates of pre-k across past studies may be attributed to differences in research design or measurement limits (e.g., ceiling effects on a measure may obscure enduring differences across groups), much of the variance in impact estimates can be attributed to substantive differences across or within studies in several observable factors: (1) characteristics of the children who receive pre-k (e.g., sex, race, parental education, income); (2) characteristics of the ECE context including other educational experiences occurring before or during the pre-k year; (3) qualities of the subsequent school and teacher environments children experience after they complete pre-k; and (4) qualities of the economic community-level environments children experience.

Our work begins by extending previous research that has shown that pre-k effects differ across groups of children that vary in race, ethnicity, sex, and mother's education status. We conceptualize these characteristics not as biological factors but as proxies for environmental experiences that children have. Many ECE programs (e.g., Abecedarian, Perry, NC Pre-K) were targeted toward and designed to support children growing up in low-income circumstances because they assumed that a low-income context fails to afford children the kind of stimulation needed to succeed in school. They expected that an "enriched" pre-k experience would compensate for the environmental disadvantage these children otherwise endure, and that children growing up in higher-income circumstances would not benefit further from pre-k because they already experience environmental enrichment. This is a hypothesized interaction effect that program developers and policymakers hope would reduce later achievement gaps and other disparities in outcomes across groups.

In this Monograph, we also test for differences in the effect of NC Pre-K based on demographic characteristics, but we extend this work by focusing on three theories that we argue make plausible, but distinct, predictions about possible patterns of heterogeneity in pre-k effects. The first theory is **dynamic complementarity** (Cunha & Heckman, 2007; and its related "sustaining environments" hypothesis, Bailey et al., 2017), which asserts that enriching educational environments occurring before, during, or after the pre-k year accelerate the impact of pre-k. If the enriching environment occurs before the pre-k year, the effect of pre-k is that "the rich get richer." Baulos and Heckman (2022) refer to this interaction as "skill begets skill." If the enriching educational environment is experienced after the pre-k year (by attending, e.g., a school that has historically high academic achievement or a high proportion of teachers with many years of experience), the advantages that children who attended pre-k enjoy over peers who did not attend pre-k will increase in magnitude. This phenomenon is observed empirically as a positive interaction effect between pre-k and another *positive* environmental exposure. In other words, children who have other environmental advantages would enjoy the largest benefits of pre-k exposure.

A contrary theory predicts a **compensatory effect**, empirically operationalized as a positive interaction effect between pre-k exposure and another environmental experience that exerts a *negative* influence on children's outcomes. Thus, this theory predicts that the impact of pre-k is most beneficial for children who experience lower-quality environments in other domains of life and, reciprocally, that children experiencing higher-quality environments will benefit less from pre-k. If those other negative influences occur before the pre-k year, exposure to pre-k acts as a "buffer" by blunting the adverse impact of early negative environmental experiences. If those negative influences occur after the pre-k year, the pre-k year serves as prophylactic "protection." Empirically, if this hypothesis holds, the advantages that children who experience pre-k have over peers who do not experience pre-k will be larger in magnitude under other negative conditions and will be diminished under other positive conditions. Economists often refer to such interaction effects as evidence for "substitution," because one positive influence may substitute for another without supplemental gain; that is, either factor alone can have positive impact, but both factors combined do not have higher impact than either factor alone.

Finally, a third hypothetical possibility is that pre-k and another environmental condition operate as positive influences without a significant interaction. In investment terms, pre-k and the other environmental experience would each have an **additive effect** on a child's later life course that neither grows nor fades out under various conditions. The effect of pre-k is to provide an additive increment to a child's outcome while the other environmental experience provides an additional but discrete additive increment.

In this Monograph, we analyze data from North Carolina's scaled-up NC Pre-K program (initially called "More at Four" when it began in 2001) to examine how various factors account for impact heterogeneity. We build on previous evaluation work for this program that leveraged the state's allocation of funding for pre-k in a natural experiment design (e.g., Dodge et al., 2017; Ladd et al., 2014). This previous work has reported the robust positive impacts of state funding for pre-k on children's reading and math achievement, grade retention, and special education placements through 8th grade (Bai et al., 2020). We begin by extending these main effect analyses to six additional cohorts, and we update the analytic models with new covariates, attempting to isolate exogenous variation in funding for the program over time in a more robust manner. Like previous work on this program (e.g., Dodge et al., 2017; Ladd et al., 2014), our work also examines the impact of funding on academic achievement (reading and mathematics test scores) and school placement decisions (grade retention and special education placement) for children who were age-eligible for the program. We also provide a new set of exploratory instrumental variables (IV) analyses that leverage a child's NC Pre-K enrollment status to estimate impacts from actual program participation given increases in county NC Pre-K funding. Finally, we conduct heterogeneity analyses across a broad set of possible moderators. Whereas previous analyses of NC Pre-K have tested for moderation based on standard demographic characteristics (e.g., Dodge et al., 2017; Ladd et al., 2014), we substantially broaden the set of potential moderators to examine how NC Pre-K funding interacts with a host of additional environmental factors.

Outline

Our Monograph Proceeds with the Following Outline

In Chapter II, we begin with a detailed overview of key findings from high-quality ECE studies to help contextualize our examination of NC Pre-K. We provide a brief review of small demonstration programs administered by researchers in highly controlled environments, because these studies have had substantial influence on ECE research and policy. We then move to a review of the effects of large-scale pre-k programs, including HS, with a focus on programs that have been evaluated through the use of rigorous research designs.

In Chapter III, we turn to explanations of fadeout or persistence and review studies that have leveraged environmental heterogeneity to understand the presence or absence of longer-term effects reported in high-quality ECE research. We present evidence for each of the three plausible patterns that hold implications for developmental theory regarding the long-term effects of early investments: (1) complementary effects; (2) compensatory effects; and (3) two main effects with no significant interaction (an additive effect).

We next discuss our data set, measures, and analytic plan in Chapter IV, before presenting results in Chapters V (main effects) and VI (environmental heterogeneity).

In Chapter V, we make several updates to previous work (Dodge et al., 2017; Ladd et al., 2014) that reported average impacts of North Carolina's pre-k investments on children. We report new main effect estimates of the impact of funding for NC Pre-K on academic achievement, special education placement, and grade retention in 5th grade. These new analyses include six additional cohorts of students who were exposed to program funding in years not considered by previous evaluations (i.e., 12 cohorts were included in Dodge et al., 2017; Ladd et al., 2014). We then examine variation in the NC Pre-K funding main effect by child and family characteristics through subpopulation analyses of race and ethnicity, birthweight status (low vs. normal/high), and maternal education (less than high school vs. high school or higher). Importantly, we report new analyses that incorporate a child's enrollment status in NC Pre-K—another factor that had not been considered in previous work using this design. In these analyses, we test the impact of enrollment in NC Pre-K on 5th-grade achievement by using our measure of NC Pre-K funding as an “instrument” for a child's enrollment. These analyses attempt to purge enrollment of selection bias by using only the variation in pre-k enrollment that is determined by NC Pre-K funding. With these models, we also test whether enrollment in the pre-k program differs across key subgroups.

In Chapter VI, we describe how we test potential sources of environmental heterogeneity in the impacts of NC Pre-K investments that focus on characteristics of children, access to alternative ECE programs, features of their subsequent schooling environments, and the economic conditions of their larger community. The environmental heterogeneity results represent the clearest tests of our hypotheses regarding complementarity, compensatory effects, and additive effects. We also examine moderation intersectionality by testing for heterogeneity in environmental moderation across each of the subpopulations of interest.

Finally, in Chapter VII, we summarize our results and integrate our findings into the broader pre-k literature. We conclude by considering how our results inform developmental theory regarding the effects of early educational experiences and how our findings can inform the continued scale-up of pre-k programs across the country.

II. Review of ECE Program Main Effects

Findings from Legacy Programs

Much of the hope surrounding early childhood interventions is fueled by promising evidence from two small evaluations of programs demonstrating the potential long-term benefits of substantial investments in early childhood environments. The Perry Preschool Program and the Abecedarian Project, which we refer to as “legacy” programs due to their long-lasting influence on the field of ECE research, were implemented 50 years ago and evaluated through RCTs with small numbers of children. Both studies indicated that high-quality early childhood interventions administered with tight control by program developers can produce immediate positive impacts on children that persist into adulthood (see Elango et al., 2016, for a review). The unique features of these programs make effects from these studies difficult to generalize to policy discussions today, a point we return to below. However, these studies act as proof points for the idea that high-quality early investments can have lasting impacts on children's lives.

The HighScope Perry Preschool Program was implemented in Michigan beginning in 1962 and evaluated by randomly assigning 123 low-income Black 3-year-old children to an intervention or control group. The intervention ran for 2 years at ages 3 and 4 and consisted of a 2.5-hr/day, in-school “active learning” curriculum, along with biweekly home visits. Participants have been followed through mid-life. Consistent with an unjust norm of denied resources for Black families in the 1960s, children in the control condition received little out-of-home support for learning. Children assigned to the Perry intervention had stronger school achievement and intelligence test scores at the end of the preschool intervention. By age 8, the impact on intelligence scores had faded to nonsignificance, though impacts on academic achievement measures persisted (Duncan et al., 2022). The children from the Perry Preschool study have been followed for decades, and evaluation analyses have indicated the program had crucial sustained impacts that were detected well into adulthood. The treated group ultimately presented with higher high school graduation rates, increased earnings, and reduced participation in criminal acts (Heckman et al., 2010; Schweinhart et al., 2005).

Much like Perry Preschool, the Carolina Abecedarian intervention was also directed toward mostly Black, low-income families in the 1970s. Beginning in infancy, the intervention lasted through kindergarten matriculation and included year-round, full-day, high-quality out-of-home care, as well as regular home visiting. Similar to the Perry Preschool evaluation, 111 children living in substantial poverty were randomly assigned to intervention or control. The control group's circumstances had improved since the 1960's era of the Perry Program: 75% attended some type of alternate ECE program before entering school (García et al., 2020), and all children were followed into adulthood. The treatment-assigned group had higher scores on reading and mathematics achievement tests that were found early and persisted over time. In adulthood, intervention-assigned individuals had higher rates of college graduation and full-time employment, as well as less involvement with crime (Campbell et al., 2012, 2001; García et al., 2020).

The long-term positive effects of these demonstration studies have been highly cited and are often used to motivate publicly funded interventions that have been scaled for delivery to large numbers of children. However, findings from the Perry Preschool and Abecedarian studies may have only limited potential for projecting the long-term impacts of the scaled-up preschool programs considered by policymakers today. The external validity of these studies remains controversial for several reasons, including the fact that counterfactual conditions from the 1960s and 1970s have changed substantially in the last 50 years. The "control" children in today's studies undoubtedly experience more support than low-income Black children received in the 1960s and 1970s. Unlike 50 years ago, today's baseline conditions include Medicaid, the Earned Income Tax Credit (EITC), Sesame Street, some Internet access, less exposure to lead, and increased access to higher-quality out-of-home childcare. The "bar" that an enhanced intervention program must exceed is higher today than it was 50 years ago. Moreover, Perry and Abecedarian were intensive and comprehensive programs tightly monitored by researchers, and population-level per-child funding allocations today are much lower than what would be required to implement these legacy programs now (Duncan et al., 2022). Large-scale pre-k programs may not employ uniform quality control measures of the level implemented in the legacy trials. On the positive side, Bruno and Iruka (2022) point out that programs today benefit from evolution in culture, growing scientific knowledge, and stronger stakeholder input. Program developers in the 1960s were bound by the context of the era and took a deficit-based, rather than strengths-based, approach. Bruno and Iruka argue that they did not attend sufficiently to family voice, and they took a "race-blind" approach to intervention design rather than being race-responsive.

Thus, to understand better how today's investments in ECE might affect children in the long-term, we must evaluate evidence that has external validity to today's cultural and policy contexts. We now turn to studies that have attempted to evaluate recent public investments in ECE operating at scale.

In 2020, total state funding for public pre-k surpassed $9 billion, and approximately one third of all 4-year-olds in the country (~1.4 million children) were enrolled in publicly-funded preschool (Friedman-Krauss et al., 2021) with wide variation across communities. The growing footprint of state-funded pre-k programs has led to substantial interest from researchers in evaluating the effects of these investments on children. Correlational research using national samples finds that increased enrollment in public pre-k has led to academic benefits for children during the early elementary years (Bassok et al., 2018; Claessens et al., 2013), and numerous nonexperimental studies by academics (e.g., Huang et al., 2012) and state-funded research agencies (e.g., Texas Education Agency, 2017) have shown positive effects of pre-k on subsequent student achievement.

However, approaches that simply compare pre-k attenders to non-attenders, even when controlling for demographic variables, suffer from internal validity concerns due to unobserved factors that may lead to selection into pre-k programs. Phillips and Pre-Kindergarten Task Force (2017) convened a diverse group of scholars to review the most rigorous evaluations to date of scaled-up public pre-k programs. They issued a consensus conclusion that most evaluations reveal a significant positive impact on academic skills measured at the conclusion of the pre-k year, when children matriculate into kindergarten. After elementary school begins, however, findings diverge. Some program effects fade completely, some fade partially, and some have sustained impact.

Concerns about fadeout have become even more pronounced in recent years (Duncan et al., 2022). Duncan and Magnuson (2013) concluded, "most early childhood education studies that have tracked children beyond the end of the program treatment find that effects on test scores fade over time" (p. 120). Bailey et al. (2017) reviewed 67 ECE studies that followed children after program completion and found an average immediate posttreatment positive effect of intervention of 0.23 standard deviations (*SD*s) that faded to 0.06 *SD*s 4 or more years later.

In the following sections, we review evidence from studies that have used rigorous methods for program evaluation, including experimental and natural-experimental designs, which also include some follow-up evaluation to detect persistence or fadeout. We focus on findings from evaluations of public pre-k programs operating in Massachusetts, Tennessee, Oklahoma, Georgia, and North Carolina. Several other rigorous evaluations have tested the effects of pre-k curricular interventions, which alter the quality of children's environmental experiences in early childhood educational programs compared with "business as usual" pre-k (e.g., Bierman et al., 2008; Clements et al., 2011; Nesbitt & Farran, 2021; Raver et al., 2011). Though we do not focus explicitly on studies examining the efficacy of specific program features (e.g., curricula, teacher credentials, etc.), we attempt to provide details regarding the salient features of each program, including information regarding the quality benchmark ratings of the program determined by the National Institute for Early Educational Research (NIEER; see Friedman-Kraus et al., 2021). It should be noted that our review of the literature is selective and primarily designed to motivate the theoretical approach and analyses that follow. Interested readers should turn to recent reviews by Duncan et al. (2022) and Cascio (2021) to find more comprehensive assessments of this literature.

Boston Public Schools Pre-Kindergarten Program

The Boston Public School Pre-Kindergarten program has been the focus of substantial research attention. It offers children full-day pre-kindergarten with empirically-supported curricula, including the *Opening the World of Learning* (OWL) literacy and language program (Ashe et al., 2009) and the *Building Blocks* (Clements & Sarama, 2007) early mathematics curriculum. Boston pre-k teachers are held to a high standard. They are required to have a bachelor's degree (and they must be on track to receive a master's degree in the next 5 years), and they are paid the same rate as other K-12 teachers in the district. Perhaps most importantly, they receive continuous coaching on how to implement pre-k curricula in the form of weekly to bi-weekly on-site support. Not surprisingly, Boston's pre-k program has regularly been seen as a “gold standard” program by the research community.

The first rigorous evaluation of the Boston program was conducted by Weiland and Yoshikawa (2013), using a birthdate-based regression discontinuity (RD) design that compares two groups of similar children: those who were eligible to participate in pre-k and those who were almost the same age but not yet eligible for participation because they barely missed the age-cutoff eligibility requirements of the program. (See another example of this design by Gormley et al., 2005.) The sample consisted of 2,018 4-year-old children, who were diverse across a number of demographic characteristics: 41% were Hispanic, 26% Black, and 50% multilingual. Findings at kindergarten entry revealed moderate-to-large positive impacts (effect sizes of .45–.62) on direct assessments of language, literacy, and mathematics, and smaller positive impacts on measures of executive functioning and emotional regulation (effect sizes ranging from .03 [n.s.] to .28).

These mostly positive and significant short-run effects on developmental outcomes are complicated by two follow-up studies that leveraged lotteries at over-subscribed programs to approximate RCTs. These studies take advantage of sites that have more demand than available slots, and prospective enrollees are essentially randomly assigned the opportunity to attend a given program based on lottery results. Although this design has the advantage of random assignment, which improves internal validity, it suffers from external validity problems because the communities are populated by relatively high numbers of parents who are motivated to register their excluded children into pre-k. It can be presumed that "control" parents who do not "win" the lottery are more likely to try to find alternative positive educational experiences for their children than would typical nonregistered parents, and these communities might have more high-quality alternatives to pre-k in response to unmet demand than would the average community.

Working within this design and using state achievement data, Weiland and colleagues (2020) found no statistically significant association between winning the lottery and math and reading test scores in 3rd grade in a sample of children with similar demographic characteristics to their prior regression-discontinuity work (i.e., 38% Hispanic, 24% Black, 25% White; 57% eligible for free/reduced lunch). Null effects were also reported for special education placement and retention. It should be noted that, almost all families who lost the lottery for the pre-k program of their choice were able to enroll in other forms of formal educational care. One significant effect was found: children who won the lottery were almost twice as likely to remain enrolled in Boston Public Schools, suggesting that families who lost the lottery may have found alternate nonpublic schools and education experiences for their children. Interestingly, recent work relying on propensity scores for balancing differences in characteristics between enrolled and nonenrolled children found some evidence suggesting that Boston pre-k impacts on achievement could be detected as late as Grade 3, though much of the initial positive effect of participation in the program still faded during the kindergarten year (Weiland et al., 2021).

Gray-Lobe et al. (2021) also leveraged lotteries from oversubscribed Boston pre-k centers to study long-term impacts. However, they relied on an earlier cohort of students, which was comprised of higher shares of Black students (41%) and had slightly fewer Hispanic students (34%) than the Weiland et al. studies. This earlier cohort of students might have also had a different set of alternate educational opportunities than the "control" children in the Weiland et al. (2020) study. (Few details are provided about counterfactual conditions in the Gray-Lobe et al. study.) Further, the lottery winners received an earlier version of the Boston pre-k program that lacked many of the high-quality features that were later introduced. Nonetheless, this earlier cohort could be followed for a longer period, allowing researchers to evaluate longer-term impacts of pre-k attendance on college-enrollment, college preparation, standardized test scores, and behavioral outcomes. They found the lottery offer of pre-k enrollment led to increased college attendance, SAT test-taking, and high school graduation, and the program decreased juvenile incarceration. As with the findings reported by Weiland et al. (2020), they found no impact on state achievement test scores in third through 10th grade.

Tennessee Voluntary Pre-K Program (TNVPK)

The TNVPK began at a small scale in 1996, before expanding over time to serve about 18,000 low-income 4-year-old children annually. To receive state funding, public school preschool classrooms must devote a minimum instructional time of 5.5 hr/day, 5 days a week, in classrooms of no more than 20 students, led by a state-licensed teacher using a state-approved curriculum. Lipsey et al. (2018) describe the pre-k environment in TN as follows: "(T)here is considerable diversity in local implementations. In this regard, the Tennessee Voluntary Pre-k Program (VPK) is not atypical of state pre-k programs generally, operating with some mandated structure based on accepted standards, but neither tightly controlled nor shaped and guided by an overarching vision widely understood and embraced throughout the state" (p. 157). The TN program has been generally rated favorably by the NIEER system. Overall, the program met an average of 8.3 of the 10 benchmarks for quality set forth by NIEER in the years between 2005 and 2021.

The TNVPK program was evaluated using an RCT by Lipsey et al. (2018). Similar to work in Boston, the RCT was run in oversubscribed sites, allowing researchers to randomly assign children the opportunity to attend the TNPVK, but with the implication that families that did not win this lottery may have been motivated to find educational opportunity elsewhere. Their study included 2,990 4-year-old children applying for admission in sites spread across the state, and lottery applicants reflected the diversity of the program (27% Black, 23% Hispanic, 49% White, 24% English language-learners). Using state-supplied administrative data, Lipsey and colleagues reported mainly null-to-adverse effects of random assignment to the program on achievement scores in 3rd grade (treatment-on-the-treated effect sizes ranged from −.13 for reading [n.s.] and −.23 for math [$p < .05$]). They also found random assignment to pre-k was associated with a higher probability of being placed in special education during the earliest years of elementary school, which may have occurred because children spent an additional year in the public school system, thereby increasing surveillance of their eligibility for special education. These initial findings were corroborated by an independent reanalysis of the study data by Watts et al. (2019). More troubling, the negative effects on achievement were shown to extend into 6th grade, and the pre-k group experienced worse behavioral outcomes for disciplinary infractions and attendance in early adolescence (Durkin et al., 2022).

The surprising negative impacts of the TNVPK RCT have led to many questions about the possible factors that could have led to unintended effects. Certainly, the program does not have the same quality standards as those employed in more recent years by Boston, yet the point estimates on 3rd grade math test scores in both studies (i.e., Lipsey et al., 2018; Weiland et al., 2020) were similar in magnitude (e.g., TOT on math in Boston: −.18, n.s.; TOT on math in TNVPK: −.23, $p < .05$). Further, Lipsey and colleagues recruited a smaller sample of children participating in the statewide lottery study to participate in more intensive data collection spanning from pre-k through Grade 3, which allowed researchers to observe how pre-k impacts unfolded over the course of elementary school. These "intensive sub-study" students were administered the Woodcock Johnson Tests of Achievement at each grade. Students randomly assigned to the TNVPK program had higher scores on five of six subtests at the end of pre-k, with effect sizes ranging from .16 to .28 (Lipsey et al., 2018). Thus, if the program was of low quality, it still produced positive initial effects on key measures of achievement. Following the fadeout pattern described by Bailey et al. (2017), these initial positive impacts faded to zero during kindergarten and Grade 1, before turning negative later in elementary school. Most impacts on Grade-2 and Grade-3 assessments were negative but not statistically significant.

Additional analyses of the TNPVK RCT data indicate that adverse impacts did not hold for every group or in all circumstances, as will be described in depth in Chapter III. But, the main effect analyses indicate that lottery-based assignment to TNVPK in over-registered public-school sites had positive initial impacts on academic skills. This impact was not sustained across early elementary school years and the effect became negative in later grades.

Tulsa Pre-K Program

The Tulsa Pre-K Program in Oklahoma has received considerable attention from Gormley and colleagues across a number of studies that have employed various natural experimental designs. Oklahoma's universal pre-k program for 4-year-old children began in 1998, and the program is regarded as having high-quality standards and services: student–teacher ratios cannot exceed 10:1, and teachers must have a bachelor's degree and must follow the state-mandated curriculum guidelines. Program quality, teacher credentials, and child participation rates have steadily increased over time. The NIEER quality ratings for the OK pre-k program look much like the TN program. Between 2003 and 2021, OK met, on average, 8.5 NIEER benchmarks.

Gormley et al. (2005) also used a birthdate cutoff RD design to examine impacts of the program on school readiness skills. The Tulsa RD design was implemented across two cohorts; approximately 40% of the sample was Black, approximately 15% was Hispanic, and approximately 35% was White; most of the children qualified for free or reduced-price lunch. Gormley and colleagues reported large effects on children's achievement at school entry, with effect sizes ranging from .38 to .79.

Because the birthdate RD compares children who just finished a year of pre-k to similarly-aged children who were about to begin a year of pre-k, the design does not allow for longitudinal analyses of pre-k effects. Thus, Gormley and his team investigated longer-term effects of the Tulsa Pre-K Program by using propensity score matching that adjusts for differences in children's backgrounds that might lead to differential selection into pre-k (Gormley et al., 2018; Hill et al., 2015). These studies have largely found that, conditional on covariates used in the matching process, children who enrolled in pre-k have some better long-term outcomes than similar children who did not enroll. For example, Hill, Gormley and Adelstein (2015) found no lasting achievement impacts for the first cohort of students but sustained impact for the second cohort on 3rd-grade math scores. By 7th grade, propensity-score adjusted estimates suggested that the program effects continued for math, honors course enrollment, and grade retention, although the magnitude of impact had reduced (Gormley et al., 2018).

Georgia's Pre-K Program

Like Oklahoma, Georgia was also an early adopter of state-funded pre-k, and this initiative was accelerated by President Obama's Preschool for All plan. Georgia's pre-k program serves 4-years-olds across the state regardless of family income, and includes a number of quality measures. As such, providers are required to use a state-approved curriculum, and lead teachers must hold at least a bachelor's degree (Peisner-Feinberg et al., 2014). Consequently, Georgia has also maintained relatively high and consistent NIEER quality ratings over time. In most years since 2005, the program has been rated as meeting 8/10 of the NIEER standards, and had an average rating of 7.9/10 between 2005 and 2021.

Initial evaluations suggested the introduction of universal pre-k in Georgia was associated with small but positive impacts on 4th-grade National Assessment of Educational Progress (NAEP) scores in the state between 1994 and 2005 (Fitzpatrick, 2008). Most students included in the analysis were either Black (ranging from 33% to 48% across years as the racial demography changed) or White (ranging from 44% to 63% across years). Further, an RD study, which included a demographically similar set of students (53% White; 38% Black), reported significant positive effects on children's language, literacy, and math skills at kindergarten entry (Peisner-Feinberg et al., 2014).

These results were further bolstered by Cascio and Schanzenbach's (2013) difference-in-differences study, which compared outcomes for children in both Georgia and Oklahoma to outcomes in the rest of the United States during the rollout of the universal pre-k programs in both states. They replicated Fitzpatrick's finding of positive impact on Grade-4 NAEP math scores for children in Georgia (with slightly larger estimates). Estimates for Grade-8 math scores were still positive, but results suggested important heterogeneity of impact across groups of children. Impact at Grade 8 was positive for children receiving free or reduced-price lunch (~0.07 *SD*s), but no impact was detected for children not receiving subsidized lunch. Because free-lunch eligibility is known to correlate with lower test performance, the pre-k experience may protect these children from poor outcomes.

North Carolina's Pre-K Program (NC Pre-K)

The NC Pre-K is the focus of the current study, and consequently, we provide extensive details regarding program features and previous empirical work below. Our description of the history of the program relies heavily on the details provided by Ladd et al. (2014), Peisner-Feinberg (2003), and Peisner-Feinberg and Schaaf (2009).

North Carolina began its statewide, low-income-targeted pre-k program in 2001 when Mike Easley took office as Governor. The program, initially titled More at Four,[2] was created with the goal of promoting the school readiness skills of disadvantaged children. Four-year-old children are eligible to receive NC Pre-K funding if their family has a gross family income at or below 75% of the state median income (SMI). Eligibility is also extended if the child has an educational or developmental delay, an identified disability, a chronic health condition, is an English language-learner, has a parent serving in the military, or is homeless.

The NC Pre-K Program has also tried to improve outcomes for the entire population, particularly for low-income children, by leveraging state funds to incentivize local pre-k programs to improve quality and serve more low-income children. From the beginning, it has operated through a mixed delivery system by providing funding to qualifying providers where parents may choose to send their children, rather than creating new classrooms in public schools, as in Boston, Tulsa, and Tennessee. In turn, pre-k children are served in a variety of settings, with approximately half served in public schools and the rest attending HS centers, private for-profit childcare centers, and nonprofit centers (Peisner-Feinberg & Schaaf, 2009). This provision meant that state funding supported eligible "slots" rather than classrooms, that many state-funded children attended the same pre-k classrooms as nonfunded children, and funded children went on to attend the same elementary-school classrooms as nonfunded children. The state program designers explicitly intended this provision to enhance "spillover" effects of state funding on nonfunded children. To promote quality, sites could not receive funding for NC Pre-K slots unless they complied with state guidelines involving classroom curricula, training and education levels for teachers and administrators, class size, and in offering additional program services (Peisner-Feinberg & Schaaf, 2009). The financial incentive for a provider to improve to meet eligibility standards was high, and on average childcare centers in North Carolina improved substantially over time. By the 2009–2010 year, NC Pre-K met all 10 of the quality benchmarks established by NIEER—representing an increase from 7 quality benchmarks met during the 2002–2003 year (Barnett et al., 2010, 2003). Between 2002 and 2021, the NC program met an average of 9.3 of the 10 benchmark standards set by NIEER. As the quality of pre-k in North Carolina increased, the proportion of 4-year-olds attending pre-k also increased. The results of this provision were that more children attended pre-k, and that all children attending pre-k (both funded and nonfunded) experienced higher-quality education. For this reason, we view the NC Pre-K program as a population-level intervention that is best evaluated by its population impact.

Over this time span, the program grew extensively. In the first year, local community organizations in NC applied to the state for funds, and 28 of these applications were approved for funding. This led to 1244 children enrolling in the program across a variety of settings during that year. By the 2003–2004 academic year, all 100 counties in the state received funding, with 628 local sites participating in the program, and nearly 11,000 children enrolled in NC Pre-K (Peisner-Feinberg & Schaaf, 2009). By 2008–2009, the program had expanded to 1,285 local sites serving over 33,000 children (Peisner-Feinberg & Schaaf, 2009).

Much of the research on the NC Pre-K program has been conducted by Peisner-Feinberg and her team at the Frank Porter Graham (FPG) Institute at the University of North Carolina, which held multiple contracts to evaluate the program across the first 19 years of the program's rollout. The initial reports (e.g., Peisner-Feinberg, 2003) provided descriptive information on classrooms and children participating in the program. Using a random sample of classrooms receiving pre-k funding, Peisner-Feinberg and Schaaf (2009) reported that participating classrooms were rated favorably on classroom observational measures of quality, such as the ECERS and CLASS. They also found that over the decade, the vast majority of centers used Creative Curriculum, with the total share of classrooms using the curriculum growing to nearly 87% by 2008–2009. These initial reports also found descriptive evidence that children participating in the program showed strong growth in language, math, general knowledge, and social skills (Peisner-Feinberg & Schaaf, 2009). Because the study did not have a comparison group, these findings do not provide rigorous evaluations of program effectiveness.

Peisner-Feinberg expanded on this work by conducting several studies that constructed comparison groups using various quasi-experimental approaches. The first study was an RD study that compared pre-k attenders and non-attenders. The RD study sample included 509 pre-k attenders and 501 non-attenders. Analyses revealed post pre-k differences favoring pre-k enrollees on all measured language and literacy skills (letter-word knowledge, phonological awareness, print knowledge) and math (applied math problems, counting), except for receptive vocabulary (Peisner-Feinberg & Schaaf, 2011). Another study evaluated all 3rd-graders in North Carolina across two cohorts and included a sample of 5,554 pre-k attenders and 200,062 nonattenders, including a sub-sample of 4,065 attenders and 96,252 nonattenders who qualified for free/reduced price lunch in 3rd grade. Among those who qualified for free/reduced price lunch, pre-k-attending children had higher 3rd-grade test scores than nonattenders (Peisner-Feinberg & Schaaf, 2010). Finally, a third study from this team used propensity scores to compare pre-k attenders to nonattenders in a sample that was 50% White and 30% Black, with approximately 29% identifying as Hispanic in terms of ethnicity, and 20% English language learners. This study found positive associations with math and executive function skills at the end of kindergarten, with largely null effects on measures of literacy and language (Peisner-Feinberg, Mokrova, et al., 2017).

More recent research from Peisner-Feinberg and colleagues (2020, 2019) used an RCT design, which leveraged a sample of children attempting to attend oversubscribed programs in two counties with large waitlists. All children participating in the RCTwere screened for eligibility for the NC Pre-K program, and children were largely drawn from economically disadvantaged households (89% with earnings below 75% of the SMI) that were demographically diverse (57% Black, 36% White; 34% Hispanic ethnicity; 42% limited English proficiency). Conclusions from this study are somewhat hampered by the fact that only 26% of the full randomized sample ($n = 582$ out of 2243) agreed to participate in the study. The pattern of results is similar to those from other RCTs (e.g., Lipsey et al., 2018), in that children randomized to the pre-k group had stronger language and literacy skills at the end of pre-k, though no significant impacts were detected on executive function, math, or behavioral skills (Peisner-Feinberg et al., 2019). However, significant main impacts faded during kindergarten, save for some persistent effects on vocabulary skills for dual-language learners (DLLs). (We return to issues of heterogeneity of program impact in Chapter III; Peisner-Feinberg et al., 2020.)

A different approach to evaluating the NC Pre-K program was taken by Ladd, Muschkin, and Dodge over the course of a series of papers (Bai et al., 2020; Dodge et al., 2017; Ladd et al., 2014), which capitalized on the staggered roll-out of NC Pre-K funding during the early years of program scale-up. Ladd and colleagues (2014) reasoned that differences between counties over time in the level of state funding provided for the program might be idiosyncratic, and therefore could be leveraged as a type of natural experiment that could identify impacts of funding on child outcomes. This assumption rested on the somewhat haphazard nature of the program roll-out, as the State Legislature determined the level of funding each year, and the Department of Public Instruction doled out this funding to counties based on imperfect estimates of child population and program capacity. Funding was also spread across Congressional districts in a way that ignored need. During the initial program years, program administrators were tasked with collaborating with the local ECE systems that were already in place, including NC's older early childhood initiative, Smart Start. Indeed, the NC Pre-K program has been primarily administered at a local level by either the public school system or the local Smart Start partnership (Peisner-Feinberg et al., 2020). Thus, funding was somewhat conditional on a county's level of potential capacity for ECE, the level of need given the population of eligible 4-year-old children, and the potential for collaboration between the local Smart Start office, HS centers, and the public school system.

The Ladd et al. (2014) research design capitalized on this variation in year-to-year funding levels using a two-way fixed effects model, which leveraged differences across years in North Carolina's 100 counties. This model uses both year and county indicator variables so that the effects of funding changes are identified by shifts within counties over time, controlling for any additional statewide unmeasured changes that might affect the outcomes. This design essentially follows the logic of a DD design with a continuous "treatment" variable. Ladd and colleagues (2014) observed that the scale-up in funding levels across counties and years appeared to be exogenous to most observed child characteristics and family residential movement. Importantly, they also found that their model was robust even with controls for levels of Smart Start funding in any given county over time. Smart Start may constitute the most plausible confound of any estimate linking NC Pre-K funding to child outcomes, given that NC Pre-K funding was often administered through already-established Smart Start partnerships.

These DD evaluations included the entire state's population of approximately one million 4-year-old state-resident children over a 13-year period. Although the demographic composition of North Carolina has shifted over the years with more Hispanic families moving into the state, the overall demographic characteristics of the sample across the study period were diverse; approximately 61% were identified as White, 30% Black, 4% Hispanic, and mothers completed an average of 12.5 years of education. Results across the papers leveraging this design have suggested that as the funding level increased for a given county-year, children's Grade-3 standardized scores in reading and math increased, and grade retention and placements into special education declined (Ladd et al., 2014; Muschkin et al., 2015). The average per-child state investment in NC Pre-K was associated with an average gain of 0.21 *SD*s in 3rd grade reading scores and 0.24 *SD*s in 3rd grade math scores. These gains occurred for the entire population of children living in funded counties, not just the ~25% of children in the state who received a funded NC Pre-K slot. The authors reasoned that this impact was so large that it must reflect positive spillover to same-cohort and same-county peer children who did not receive funded slots. Indeed, the design of the NC Pre-K program meant that children funded by the program would go to preschool with children who were not eligible for program funding. Peisner-Feinberg and Schaaf (2009) reported that in the early years of the program, approximately 30% of children in preschool classrooms of participating centers were not funded by NC Pre-K, and this proportion was about 20% in 2008–2009. Spillover could have affected these students because funding for the program generated preschool center quality improvements. Further, these students (both state-funded and privately paying) entered kindergarten classrooms with nonattending peers, which could have also led to opportunities for spillover through peer modeling effects and effects on teachers who could teach to a higher level.

Follow-up in Grades 4 and 5 indicated that the impact of NC Pre-K did not diminish as children aged: on the contrary, effect sizes grew across Grades 3, 4, and 5 (Dodge et al., 2017). Analyses of outcomes in Grades 6, 7, and 8 again suggested that impact continued to increase as children advanced in age. The cumulative effect of average NC Pre-K funding to which a 4-year-old child was exposed eventuated in about 5 months of learning in math and reading by the time a child completed 8th grade (Bai et al., 2020).

Although our current work focuses on the impacts of a state-funded pre-k program, it is also important to review findings from the body of research on the federal HS program. HS is the nation's single largest and longest-standing investment in early childhood intervention for low-income children. Because it incorporates a year-long pre-k experience, we include it here, but HS also allows enrollment as early as age 3 and includes additional services of nutritious meals, medical and dental services, and family-level supports such as family success planning, referrals to community agencies for services, and parent engagement in program governance. Despite the advancing scale-up of state-funded pre-k programs, HS has retained a large footprint in early educational programming across the country, serving nearly 950,000 children annually (Administration for Children and Families, 2021). Indeed, many state programs directly intersect with HS, either by braiding funding (as often happens with NC Pre-K), or by providing successive experiences to children who enroll in both programs over time (Chaudry & Datta, 2017; Jenkins et al., 2016).

Most evaluations using natural experimental designs have found that HS participants have better long-term outcomes compared with similar nonparticipants (Deming, 2009; Garces et al., 2002; Ludwig & Miller, 2007). Studies that leveraged within-family comparisons of siblings have often found long-run positive impacts on adult outcomes, including educational attainment and early career earnings (Garces et al., 2002). Deming's (2009) sibling fixed effects analysis found a positive impact of HS attendance on an index of adult outcomes, which included high school graduation, college attendance, criminal activity, teen parenthood, health status, and an indicator of “idleness” (i.e., not a student and not working). However, a recent replication and extension of this analysis found that these apparent benefits were not conferred to later cohorts of HS participants, with nonsignificant effects reported for college graduation and earnings, and negative effects reported for later cohorts on the index of adult outcomes (Pages et al., 2020). These differences in impact across cohorts could be due to dilution of the quality of HS across years, improvements in the experiences afforded to “control” children in later years, or changes in economic conditions that affected impact in an unknown way.

These findings are brought into question by the results of the national Head Start Impact Study, an RCT that included two cohorts of about 4,400 children attempting to enroll in over 300 HS sites. Much like the studies employed in Boston (Weiland et al., 2020) and Tennessee (Lipsey et al., 2018), evaluators leveraged over-subscription to randomly assign children the opportunity to enroll in HS. Like the other pre-k RCTs reviewed above, the study was restricted to a subset of centers that had greater demand for enrollment than the number of allocated slots, although the sample was constructed to be nationally representative of HS programs. At the end of the HS year, 4-year-old children who had won the lottery performed better than controls on early language and literacy skills (but not math or social-emotional skills), with an average effect size of about 0.20 *SD*s (Puma et al., 2010). These gains were entirely gone by the end of kindergarten, and the null effects, which were also detected for measures of math achievement and behavioral functioning, continued through 3rd grade (Puma et al., 2012). No further follow-up studies have been conducted.

Global Context

Although the United States led the world in creating access to public school at age 7 and older, it lags behind much of the developed world in granting access to ECE, particularly pre-k. Compulsory participation in school in the United States may start as late as age 7, whereas some nations in the Organization of Economic Cooperation and Development (OECD, 2022), including Colombia, Greece, and the Netherlands, have lowered the age at which formal schooling starts to 1 year before primary school. Compulsory schooling is universal at age 4 in Costa Rica, Luxembourg, and Switzerland, and as young as age 3 in France, Israel, and Mexico.

Because the policy trend leans toward universal access in the global context, few randomized trials have been conducted in other nations. A structured literature review by Brown et al. (2022) of 26 sources studying preschool in OECD nations indicates a growing body of descriptive evidence that participation in preschool and policies to enable access to preschool are associated with positive medium- and long-term outcomes in some nations, including acquisition of cognitive and noncognitive skills in elementary school (Europe), better high school attainment (New Zealand), and positive employment outcomes in the labor market (Norway). The spottiness of these findings implies that favorable outcomes were not always observed. Duncan et al. (2022) provide a comprehensive review of preschool program impacts drawing from international studies. They find a similarly mixed set of findings for preschool program impacts both immediately following the program and later into childhood and adulthood.

Summary

Over the past several decades, investments in state-funded pre-k programs have substantially increased, and research attention has followed. Overall, research studies using rigorous designs have tended to find positive short-term impacts of enrollment in state-funded pre-k, especially on academic skills. However, no clear pattern emerges for longer-term findings. In the next chapter, we consider whether differences in child and environmental characteristics may explain heterogeneity in the longer-run impacts of pre-k.

III. Heterogeneity in the Long-Term Effects of ECE programs

Theories to Explain Heterogeneity

The studies reviewed in Chapter II suggest that most high-quality investments in children's early educational experiences yield positive benefits immediately following the end of the program (see Phillips & Pre-Kindergarten Task Force, 2017), but the evidence is less clear regarding the long-term outcomes of public pre-k programs. Indeed, several notable evaluations have produced evidence of longer-term effects (e.g., Dodge et al., 2017; Gormley et al., 2018), whereas others have found evidence of fadeout (e.g., Lipsey et al., 2018).

The mixed evidence has led to several recent reviews attempting to provide theoretical explanations for the divergent findings across contexts, most of which rest on the premise that the effects of pre-k exposure vary systematically as a function of other environmental factors (e.g., Abenavoli, 2019; Bailey, Duncan, et al., 2020; Bailey et al., 2017). When these factors vary across (rather than within) studies, the findings could appear contradictory. Some theorists hypothesize that pre-k exposure moderates the effect of other factors, whereas other theorists frame the phenomenon as other factors moderating the effect of pre-k exposure. Both framings hypothesize an empirical interaction effect between pre-k exposure and another factor. The primary goal of this Monograph is to understand possible interaction effects involving pre-k exposure.

Dynamic Complementarity

Bailey, Duncan, and colleagues (2020) discuss two types of environmental processes that could influence the persistence or fadeout of early intervention effects. First, early interventions could interact with other child environmental experiences in a complementary fashion. This process, often referred to as “dynamic complementarity” (Johnson & Jackson, 2019) predicts that impacts from early interventions will grow in the presence of later positive educational experiences that catalyze, synergize, and enhance the impact of pre-k. These later positive experiences help a child build on the skills and capacities gained during the pre-k intervention (see intervention example in Clements et al., 2013; Cunha & Heckman, 2007). In statistical terms, this theory predicts that two high-quality environments will produce a positive interaction effect if they each provide positive experiences that are complementary (rather than redundant) to one another, and the benefit from these positive experiences will compound (i.e., increase, accelerate, exacerbate) over time. Indeed, there is a widespread belief among education stakeholders that early investments in children's learning and development must be followed-up by high-quality schooling in order to be productive in the long term (Stipek et al., 2017).

We examine the theory of dynamic complementarity by testing the interaction effect between pre-k exposure and other variables that have a significant positive relation to child outcomes. Evidence for dynamic complementarity would be found if the effect of pre-k exposure is stronger at the positive end of the continuum of the other variable (and, reciprocally, if the effect of the other variable is stronger at the higher end of pre-k exposure). A positive interaction effect indicates a synergy between two variables that independently have positive effects, such that the effect of each variable accelerates or increases nonlinearly in the presence of the other variable.

Note that we define a continuous variable as having a positive effect if, empirically, one end of its continuum has a more positive effect on a child's outcome than the other end, without regard to the label or manner of operationalizing the variable. For example, the variable "federal funding level" might seem like a positive variable, but empirically it is often related to child outcomes in a negative direction in observational data, presumably because more federal resources get allocated to higher-need schools. In this case, dynamic complementarity would be identified if the positive effect of pre-k exposure was found to be stronger in contexts where allocation of federal resources based on local need was low. Of course, this also implies that the positive effect of pre-k exposure would additionally diminish when followed by poorer quality schools, as indicated by higher levels of federal resources. The essence of dynamic complementarity is that "the rich get richer," "skills beget skills." The disparity between those with initial positive effects and those with initial negative effects widens under other conditions of positive influence.

Compensatory Effects

Bailey, Duncan, and colleagues (2020) also suggest a different theory, one through which early interventions interact with other environmental experiences through a process they call substitutability. With dynamic substitutability, two environmental experiences that provide similar benefits to children may be thought to be interchangeable. Thus, if the two experiences are substitutes for one another, a child may benefit from having environmental experience "A" or "B," but no additional benefit would be conferred from having both "A" and "B" in combination. Instead, the unique effects of receiving "A" and "B" would diminish when they were experienced in combination, since the environments offer apparently redundant benefits for child development. This can be conceptualized if we imagine two SAT tutoring programs that offer nearly identical training and skill development. Both programs could have positive and significant effects on SAT performance when considered individually, but when a given child enrolls in both programs, the benefits would diminish due to their redundancy. Empirically, we would observe two positive main effects for the programs on SAT scores, with a negative interaction capturing the diminishing benefits of receiving the programs together.

If we consider how this theory might apply to broader environmental experiences, one expects that the benefits from exposure to pre-k would be largest for children who had come from contexts of disadvantage, and consequently, had less access to other high-quality educational experiences that might offer similar benefits to pre-k. In statistical terms, this line of reasoning would predict that a measure of pre-k exposure will have a positive interaction when crossed with another environment that carries a negative influence on a child's developmental trajectory (e.g., harsh or unsupportive parenting). Conversely, one would also predict that this same preschool exposure would have a negative interaction when crossed with an environment that may provide a similar benefit to pre-k (e.g., access to instruction and materials that support academic skill development in the home). Thus, this theory suggests that enrollment in pre-k may be most beneficial to children who encounter lower-quality educational environments in other domains, as the marginal benefit of pre-k would be strongest for these children.

As a practical matter, however, two experiences may not be qualitatively "interchangeable" even if they provide measurably redundant effects on children. By definition, early education is not a substitute for later education; the two experiences occur at different periods of development. And pre-k is not a substitute for being born into adversity. Yet in instances where *positive* features of later experiences produce a *negative* interaction coefficient with pre-k, this finding still implies that children who experience both enriching environments derive less benefit than the sum of the two independent effects. For the remainder of the Monograph, we apply the term "compensatory effect" here because the effect of exposure to pre-k compensates for the ill effects of an adverse environmental experience.

Additive Effects

Finally, a third theory would predict that two experiences exert independent effects on child outcomes, but they do not interact in any meaningful way. Each experience is an investment in a child's outcomes, and these investments accumulate independently of each other. It is possible that a year of pre-k increases a child's long-term outcomes without regard to other positive or negative experiences that occur before or after the pre-k year. Thus, pre-k could have an equally positive effect for children from advantaged and disadvantaged environments. In this case, the two variables measuring environmental experiences produce incremental effects on the outcome in an additive (i.e., independent) way. In this Monograph, we use the term "additive effects" to describe a situation in which two experiences each exert an influence on a given developmental outcome, but their influence appears to be independent, and no interaction is observed.

Summary of Hypotheses

These various models of environmental interactions are depicted in Figure 1. In each of the panels we consider how exposure to high-quality pre-k might interact with subsequent exposure to a high-quality elementary school, illustrating the three key theories examined here: complementarity, compensatory effects, and additive effects. These figures assume that exposure to pre-k and a high-quality elementary school each exert positive

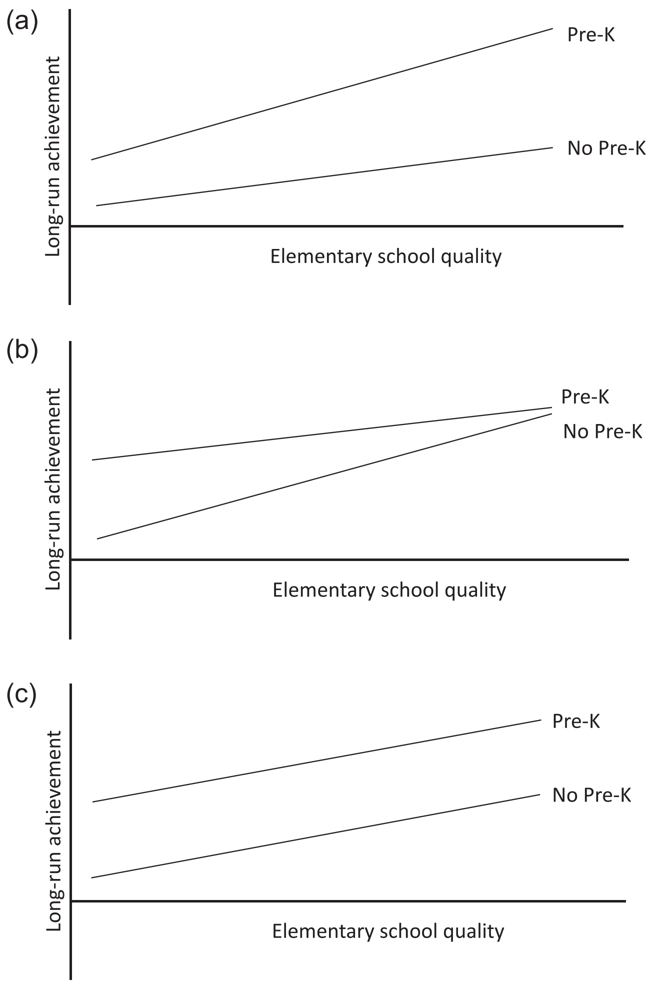


FIGURE 1.—Theoretical models of interaction effects between pre-K and later school environments. (a) Dynamic complementarity, (b) compensatory effects, and (c) additive effects.

main effects on a given child's achievement, though these same theories can also be tested in situations in which environments exert negative effects on child outcomes.

Figure 1a depicts dynamic complementarity, in which the marginal effect of exposure to pre-k is larger when children attend a high-quality school after pre-k. Note that in Figure 1a, children who did not attend pre-k still benefit from high-quality elementary school, but children who enroll in pre-k benefit more. Figure 1b depicts compensatory benefits, in which the marginal effect of exposure to pre-k is increased in the presence of lower quality subsequent schooling. In Figure 1b, both pre-k attenders and nonattenders benefit from high quality subsequent schooling, but children who attended pre-k benefit less from the later positive schooling environment. Likewise, pre-k has the strongest effect on those children who experience low-quality schooling. Finally, Figure 1c depicts two independent main effects without interaction, in which the effects of exposure to pre-k are maintained without alteration under conditions of another positive experience, suggesting a pattern of additive effects.

Evidence for Complementarity, Compensatory Effects, and Additive Effects

Traces of complementarity, compensation, and additive effects can be seen in examinations of impact heterogeneity across the ECE evaluation literature. Recent discussions regarding the prevalence of fadeout in early intervention studies have centered on the hypothesis that early intervention impacts will last only when, following ECE exposure, children encounter positive environments that sustain their prior growth. This expectation, often termed the "sustaining environments hypothesis" (see Bailey et al., 2017), would be observed in statistical analyses by a positive interaction effect, and is clearly related to the theory of dynamic complementarity. The hypothesis posits that early intervention effects will persist only when children encounter environments that specifically build upon the gains made during a given earlier educational intervention. Because children who did not experience early intervention did not acquire prerequisite skills for later skill building, they fall further behind.

The "sustaining environments hypothesis" is borne out of skill-building theories that view pre-k as a mechanism for boosting children's early skill development. Most theories of cognitive competence in math and literacy presuppose a sequence of increasingly complicated skills that build on earlier skills (Taie, 2014). This concept has been formalized in economic models of acquisition of human capital (Cunha et al., 2006). Indeed, public pre-k programs often target early academic skills, such as the fundamental building blocks of early reading, language, and mathematics. This focus on early academic skill-building has been supported by correlational research that has shown that early academic skills strongly predict achievement much later in school (Duncan et al., 2007; Watts et al., 2014), as well as labor market outcomes (Ritchie & Bates, 2013), and are considered markers of pre-k effectiveness (e.g., Gormley et al., 2005; Lipsey et al., 2018).

As others have pointed out (e.g., Bailey, Jenkins, et al., 2020; Clements et al., 2013; Stipek et al., 2017), we are most likely to observe complementary environmental interactions consistent with the "sustaining environments hypothesis" in the case of publicly funded pre-k when children enter into elementary school classrooms that have curricula and instruction designed to build on the skills gained during pre-k. This concept is known as "alignment." For example, in the case of early mathematics learning, it seems obvious that children who have mastered basic counting in pre-k will gain more from instruction in kindergarten that moves toward operations (e.g., adding and subtracting) rather than instruction that repeats principles of basic counting (Engel et al., 2012). Thus, if pre-k generates long-term impacts on academic outcomes due to boosts in children's early skill building in mathematics and reading, then pre-k would likely produce complementary interactions with later environments that provide more advanced instruction carefully aligned with the curricula encountered during preschool.

Perhaps the most rigorous study of the synergistic impact of sequenced instruction comes from a cluster RCT in which assignment to the Building Blocks pre-k math intervention was coupled with enhanced math instruction in kindergarten and Grade 1 ("follow-through"). A scale-up evaluation in Buffalo and Boston pre-k programs found that the curriculum intervention produced large impacts on children's mathematics achievement at the end of pre-k (Clements et al., 2011), and that this impact faded by the middle of elementary school (Bailey et al., 2018). However, among children assigned to the Building Blocks pre-k curriculum, those randomly assigned to the enhanced follow-through group, which provided kindergarten and Grade-1 teachers additional professional development and information on the math skills gained during pre-k, performed better at Grade 1 than did children who received only the initial curriculum intervention (Clements et al., 2013). Still, impacts for the follow-through group faded during elementary school once the follow-through treatment ended. Expanding this line of research, studies have begun to test the hypothesis that pre-k effects will persist only as long as the later school curriculum is aligned with the pre-k curriculum (Harding et al., 2020; McCormick et al., 2020).

It is less clear if we should expect complementary interactions when the features of a child's subsequent environments are operationalized in a more general way. Measures of school quality, home environment, or neighborhood disadvantage likely capture broader elements of a child's developmental experiences. If these more general environmental measures do not directly correlate with better curricular sequencing, then they may not capture environmental experiences that would lead to complementary effects. Moreover, if pre-k generates long-term impacts on children's outcomes through paths not directly dependent on academic skill building, then expectations for positive interactions between pre-k experiences and later instructional quality may also be unwarranted. To illustrate one such pathway, pre-k may produce long-term impacts on developmental outcomes because pre-k provides a safe and structured early care environment that produces healthier self-regulation skills and social-emotional functioning (e.g., Raver et al., 2008). If pre-k effects are primarily channeled through social-emotional development, then it is not clear if such impacts would be enhanced or diminished if children encounter higher quality instructional environments after preschool.

Indeed, the prevalent expectation that pre-k impacts will persist only if children encounter high-quality educational environments after pre-k is seemingly at odds with evidence that pre-k impacts are often largest for children who face systemic disadvantages, often due to racial marginalization and insufficient access to resources stemming from poverty (see review in Elango et al., 2016). Although the environmental disadvantages experienced by children from marginalized groups are often conceptualized as “baseline” characteristics in pre-k evaluation studies, there is little reason to think that the influence of such disadvantages ends when the pre-k year begins. Instead, the systemic disadvantages faced by children from impoverished families or minoritized groups are likely felt before, during, and after the pre-k year. Thus, findings that pre-k effects are often largest for marginalized groups are more consistent with the theory of compensatory effects than dynamic complementarity.

For example, Bitler et al.'s (2014) distributional analysis of HS found that impacts were largest for children at the bottom of the skills distribution. This analysis also revealed that HS effects were most persistent for Hispanic children (see also Kose, 2021), and analyses of public pre-k programs have also found larger impacts for Hispanic children, DLLs, and children eligible for free/reduced price lunch (e.g., Peisner-Feinberg, Garwood, et al., 2017; Peisner-Feinberg, Mokrova, et al., 2017; Weiland & Yoshikawa, 2013). Perhaps similarly, analyses of the Building Blocks pre-k curricula intervention also found that impacts were larger for African American children (Clements et al., 2011; Schenke et al., 2017), and correlational evidence on pre-k attendance using the ECLS-K revealed that preschool impacts were most likely to persist for Black students (Bassok et al., 2018; but see contradictory evidence in Currie & Thomas, 2000).

If pre-k effects are largest for students who have been systematically discriminated against or who are from more economically disadvantaged environments, this pattern would dovetail with long-standing developmental theory predicting that high-quality early interventions can serve as a buffer against other negative environmental experiences (Ramey & Ramey, 1998). This buffering hypothesis, which views preschool as a potentially compensatory experience, would also predict that preschool will produce positive interaction terms with other negatively valanced measures of children's environmental experiences, occurring either before or after pre-k. For clarity,

Figure 2 provides two graphical depictions of the "buffering" hypothesis during early childhood. Figure 2a graphs pre-k effects across a negative environmental exposure, such as the experience of poverty during early childhood. Here, we see that although poverty hurts the achievement of both pre-k attenders and nonattenders, enrollment in pre-k "buffers" against the negative effect of poverty, representing the long-standing motivation for publicly-funded preschool as "compensatory." In an empirical examination, one would find a negative effect of poverty exposure on achievement, but a positive interaction between poverty exposure and pre-k, indicating that pre-k effects grew larger for children exposed to deeper levels of poverty.

Yet, changing the valence of the *x*-axis variable (i.e., moderator) implies a qualitatively identical prediction of the hypothesis. In Figure 2b, pre-k exposure is graphed against early childhood family income (i.e., positively-valanced poverty). Here, we would observe positive main effects for both family income and pre-k exposure, yet the interaction between family income and pre-k would be negative, indicating that pre-k effects diminished as family income levels rose. Both figures demonstrate the "buffering" hypothesis prediction that pre-k effects will be largest for children with lower levels of family income. In essence, both figures present two sides of the same

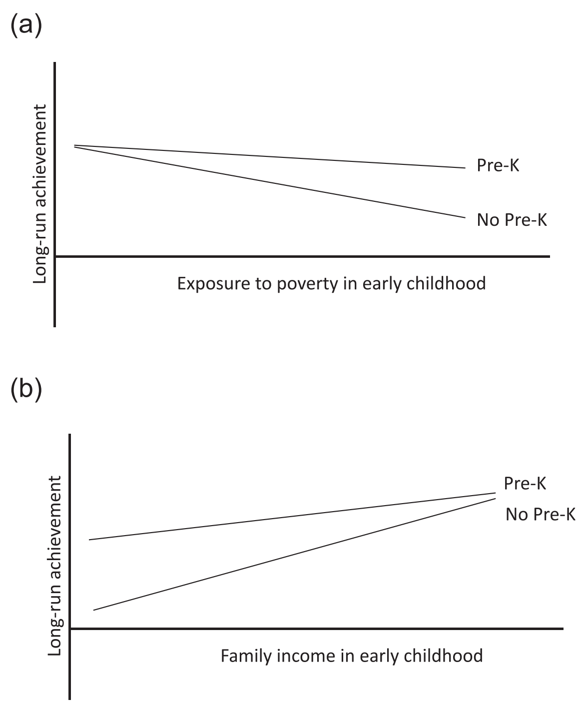


Figure 2.—Theoretical models of interaction effects between pre-K and early childhood environments. (a) Compensatory (Buffering) effects and (b) Compensatory (Buffering) effects.

coin, illustrating that compensatory effects can be observed for either "positive" or "negative" environmental experiences.

If early interventions have the largest effects for children who enter pre-k from more disadvantaged environments, might it also be the case that pre-k effects are largest for children who will *later* encounter lower quality educational environments? This logic would suggest that attending pre-k may provide a buffer against negative environmental experiences, whether such experiences are encountered before or after pre-k, consistent with the compensatory effects hypothesis. Indeed, Abenavoli (2019) described how high-quality ECE may promote resilience among children in subsequent school contexts with *lower* educational investments, evidence that higher-quality early childcare and education can buffer children from the negative impacts of lower family economic resources, including low family income (McCartney et al., 2007). Importantly, this possibility does not suggest that we send ECE attendees to low-quality schools, but rather suggests that "ECE attendees in low-quality schools are better off relative to ECE nonattendees also in low-quality schools, while the skills of ECE attendees and nonattendees converge in high-quality schools" (Abenavoli, 2019, p. 1115).

Identifying Moderator Variables

Our Monograph examines these three theoretical predictions regarding pre-k effect heterogeneity. We sought to test these processes across a wide array of possible sources of environmental heterogeneity occurring before, during, and after the pre-k year, providing a comprehensive set of potential moderators to test for the prevalence of complementarity, compensatory effects, or additive effects. This approach follows calls from numerous scholars to explore for whom, and under what conditions, ECE programs produce the best outcomes (Fisher et al., 2020; Schindler et al., 2019). The environmental experiences are measured before, during, and after pre-k, following the framework from the Phillips and Pre-Kindergarten Task Force (2017) consensus statement. Within each class of moderators, it is possible to find evidence in support of each of the hypothesized relations between NC Pre-K and a moderator (e.g., elementary school achievement). We will explain our findings using the terms from this theoretical framework—complementarity, compensatory effects, and additive effects throughout the study. From among the thousands of factors that could be tested as moderators, we focus on those that, in theory, could present a particular challenge or opportunity for pre-k to have an impact because they are substantively related to pre-k education.

In the following sections, we first briefly review *child and family individual factors* that could be understood as moderators of NC Pre-K funding effects, focusing on factors that might be interpreted as indicators of early environmental experiences (e.g., SES). Next, we turn to several classes of measures meant to explicitly capture environmental heterogeneity. We discuss how pre-k impacts could differ based on access to *alternative early childhood education programs*. Then, we examine how pre-k effects may interact with *subsequent schooling environments*. Finally, we consider how pre-k effects may depend on children's *subsequent economic environment*.

We acknowledge that there exist countless possible experiences and contextual factors that one could hypothesize to moderate long-run pre-k effects, and so our review of factors is incomplete. Consequently, we review here a stylized set of environmental and contextual factors that may moderate pre-k effects, are especially relevant for the political economy of the NC program, and take advantage of the unique affordances provided by the multiple linked sources of data we bring to bear in our study.

Characteristics of Children and Families

Consistent with the theory and empirical work on the differential and compensatory effects of ECE, several characteristics of families and children might predict who will benefit most from ECE programs. These factors represent individual strengths, but also systemic environmental disadvantage and disenfranchisement, vulnerabilities or predictors of challenges in later school settings (McCartney et al., 2007; Vandell, 2004). Although family income is broadly understood to be a key moderator of pre-k impacts (Bustamante et al., 2022; Gormley & Gayer, 2005), the NC Pre-K program was available primarily to low-income children (with some exceptions for children with special needs or in military families), so we do not examine income moderation in our study. In selecting among the thousands of possible variables, we focused on the realistic scope of what a year of pre-k educational experience could accomplish and which children might be most responsive to this type of intervention. We hypothesized that some children may, sadly, be so biologically impaired that the brief, nonintensive educational enrichment of center-based preschool may not meaningfully alter their learning trajectory, and that other children may be born with such strong environmental advantage that educational enrichment offers no added value.

In the sections below, we review moderation findings from other large-scale pre-k programs that also assessed the personal and family characteristics that we consider in the current monograph. Because interaction tests tend to be noisier, we do not note every interaction test from each state pre-k study reviewed above. Rather, we note broad patterns of heterogeneity that appear across studies.

Parental Education

Parents' education level influences parents' child development knowledge, time spent with children, and the extent to which a home environment is stimulating and engaging for young children (Davis-Kean et al., 2021; Guryan et al., 2008; Harding et al., 2015; NICHD Early Child Care Research

Network, 2001). As such, ECE programs are positioned to compensate in contexts where family educational inputs are constrained by limited parental education level. Indeed, HS impacts were strongest for children of parents with less than a high school education (Walters, 2015) and in home environments with low preacademic stimulation (Miller et al., 2014). Previous work on the NC Pre-K study, using the natural experimental design described above, has found that impacts were large for children whose mothers had lower levels of education (e.g., Ladd et al., 2014). In our study, we measure mother's education level (operationalized as having or not having a high school diploma) as a possible moderator of NC Pre-K impact.

Child Birthweight

Low birthweight is widely considered to be a risk factor for healthy development, with well-documented negative effects on short and long-run health, cognition, and attainment (Currie & Almond, 2011). The Infant Health and Development Program provided home visits and center-based preschool to low birthweight and premature infants during the first 3 years of life. The experimental evaluation of the intervention substantially improved children's cognition and behavior by age 3, with benefits sustained through age 5 and young adulthood (age 18) for children in the “heavier” low birthweight strata (Brooks-Gunn et al., 1993; Brooks-Gunn et al., 1994; McCormick et al., 2006). We examine whether NC Pre-K availability in the community is differentially effective for children who are low birthweight.

Sex

Several studies have found differential impacts of ECE interventions by the sex of the child, with most effects favoring boys (Elango et al., 2016; Magnuson et al., 2016). For example, Hill et al. (2015) found positive impact only for boys and not girls in a study of Tulsa Pre-K program's second cohort. However, other longitudinal work on Tulsa's program has found mixed results with some outcomes suggesting girls benefitted as much or more than boys when participating in the program (Gormley et al., 2018). Gray-Lobe et al. (2021) found that effects of the Boston Public Schools pre-k program on college enrollment, SAT-taking, and disciplinary outcomes were larger for boys than for girls. Similarly, previous work on NC Pre-K has also found that effects were largest for boys (Muschkin et al., 2020). In the current study, we test whether male or female students benefit more from NC Pre-K availability.

Race and Ethnicity

Numerous evaluations of ECE programs find differential impacts for children of color, especially Black and Hispanic students (e.g., Deming, 2009; Gormley & Gayer, 2005; Love et al., 2013; Puma et al., 2010). Some studies report more rapid fadeout of preschool impact among Black children, and others find sustained effects (Currie & Thomas, 2000; Gormley et al., 2018; Love et al., 2013). Weiland and Yoshikawa (2013) reported the strongest positive post-treatment impact of the Boston Pre-K program for Hispanic children, as did Gormley and colleagues (2008) in Tulsa Pre-K. Findings of stronger effects for Hispanic students may be related to language development, as other rigorous studies have found that DLLs may be among the strongest beneficiaries of pre-k participation (e.g., Peisner-Feinberg et al., 2020). We did not have access to a measure of dual-language learning in our study, but were aware that a high proportion of Hispanic students in North Carolina were DLLs. Not all studies have reported larger effects for Black and Hispanic students; some rigorous studies (e.g., Gray-Lobe et al., 2021; Gormley et al., 2005) found that results were largely consistent across racial groups. We examine whether the impact of NC Pre-K funding was moderated by a child's racial identity.

Overall, there is heterogeneity in the impact of pre-k attendance across characteristics of children and their families, but there is substantial variation in findings.

Access to Alternative ECE Programs

Next, we address whether NC Pre-K program effects are moderated by exposure to other community-level educational opportunities that children could have experienced before or during the pre-k year—namely, community availability of other early learning programs including Early Head Start (EHS) and Head Start (HS). Access and availability of other early learning programs is particularly relevant to the NC Pre-K context because of the community-level program funding allocations, mixed-delivery implementation, and population-level impact estimates derived from our natural experimental approach. As such, understanding how funding for pre-k might interact with the availability of other early childhood supports in a given community can help us understand better whether the program provides complementary or compensatory benefits to children.

Because ECE policy has evolved at the federal, state, and local levels without a coherent structure, there exists a complex patchwork of different programs, bureaucratic bodies, and funding streams for pre-k (Jenkins & Henry, 2016; National Academies of Sciences, Engineering, and Medicine, 2018). Many pre-k programs, like North Carolina's, offer care through a mixed market of private nonprofit, for-profit, and government organizations, as well as public schools. In turn, many early learning providers “blend” these multiple funding sources, whereby children in the same center or classroom may be funded by different sources (Chaudry & Datta, 2017; Duer & Jenkins, 2022; Pianta et al., 2009). Therefore, 4-year-old children enrolled in a given ECE center may see their preschool experience simultaneously influenced by NC Pre-K funding and the availability of other ECE programs (e.g., HS).

Hypotheses about the direction of impact of these co-occurring programs are mixed. One hypothesis is of synergy (e.g., complementarity), such that concurrent availability of multiple ECE programs accelerates the impact of each program. This pattern might occur if children's ECE experiences qualitatively improve when they receive funding from both NC Pre-K and other sources simultaneously. Such complementarity could also be expected if the availability of EHS (which is targeted toward children before age 3) is followed by NC Pre-K, which may provide nonredundant instruction in successive years. Similarly, NC Pre-K funding may also be more effective in communities with stronger pre-existing ECE infrastructure. If this were the case, communities that had more HS investment at the time of NC Pre-K rollout may have been more "shovel ready," and in a better position to take advantage of the early state investments in pre-k. In all of these cases, we would expect positive interactions between measures of NC Pre-K funding exposure and access to alternative forms of ECE in the community.

Alternatively, community availability of other forms of ECE could act in compensatory ways with NC Pre-K investments. With this scenario, participation in a single program brings a positive benefit, but the net gain is capped and benefits are not enhanced by the availability of multiple programs. Indeed, much of the RCT work on ECE has suggested that pre-k impacts are often largest for children who would not have had access to other forms of center-based care. Analyses of the Head Start Impact Study have found that effects were largest for children who, in the absence of treatment, would have stayed home (Feller et al., 2016; Kline & Walters, 2016; Zhai et al., 2014). These findings suggest that similar forms of center-based care likely act as substitutes for one another.

Outside of the RCT evidence on counterfactual care arrangements, few studies have examined interactions between various forms of ECE investments. The Ladd et al. (2014) work on NC Pre-K also evaluated the rollout of North Carolina's other signature statewide program for early childhood, Smart Start. The Smart Start program began in 1993 and provides support to counties to improve the early childhood system of care from birth through age 5. Thus, the Smart Start program includes investments for health, family support, and childcare. The authors found unique impacts of each program that essentially supplemented each other in an additive effect, but they did not directly test interactions between NC Pre-K funding and funding for Smart Start.

HS is a particularly important program to study in tandem with state pre-k, not only because both programs represent the most significant public investments in ECE in the United States, but also because most state pre-k programs overlap with HS in at least some service goals, and both programs prioritize the same groups of children. However, relatively few studies have examined the dynamic interaction between HS and pre-k. Bassok and colleagues (2012, 2014) have pioneered work in this area. They found that HS shifted to serve more 3-year-olds in response to pre-k expansion, but they noted other shifts in public and private-based childcare supply depended on whether state pre-k slots were funded in public schools versus the private market (Bassok et al., 2014, 2016). Researchers and policy-makers still debate the most productive ways that HS and pre-k could work together (Connors-Tadros, 2019; Gilliam & Ripple, 2004) and there are many attempts to build collaboration at state and local levels through HS collaboration offices and program liaisons (Child Care & Early Education Research Connections, 2011).

Like HS, the NC Pre-K program targeted children from low-income families. The state also imposed strong quality standards on NC Pre-K providers, and these providers were actively collaborating with HS providers across the state during the program's initial conception and rollout (Cobb, 2009). HS programs were operating in nearly all NC counties during the study period. At the start of NC Pre-K in the 2001–2002 program year, approximately 10% of NC Pre-K-enrolled children attended HS programs, growing to 20% by the 2008–2009 program year (Peisner-Feinberg & Schaaf, 2009). Furthermore, EHS serves low-income families from the prenatal period through age 3, offering even earlier intervention services, which could be additive or synergistic with HS and NC Pre-K. Given HS's quality standards, historical integration in communities, and the comprehensive services offered beyond that of pre-k, strong potential exists for synergies between HS and pre-k to benefit families and children.

In this monograph, we seek to understand how the availability of alternative childcare programs might affect long-term impacts for NC Pre-K. We examine how the presence of other state-funded early childhood initiatives (i.e., Smart Start) and federally-funded ECE programs (i.e., EHS; HS) moderate the long-term impacts of the pre-k experience.

School Experiences After Pre-K

As described in our framework above, there has been substantial investigation and theorizing about environmental moderators subsequent to pre-k intervention. We review here some of the specific constructs in this area and highlight those we examine in our study.

Studies of how subsequent elementary-school contexts differentiate the long-term effects of pre-k have yielded generally mixed findings (see review by Bailey et al., 2020). In one of the first studies of these contextual effects, Currie and Thomas (2000) found that the benefits of HS diminish for Black children who attend poor-quality schools. This is a dynamic complementarity effect because the effects of HS are strongest for those children who later attend higher-quality schools. Johnson and Jackson (2019) have one of the more compelling studies on this topic, as they leveraged exogenous variation in the rollout of HS along with exogenously mandated funding increases for public schools to examine dynamic complementarity. They followed 15,000 children in HS programs through adulthood, and found that plausibly exogenous investments in HS had a larger impact when children subsequently attended K through 12 schools that were relatively well funded.

This interaction was a complementary effect, such that the combined impact of HS and subsequent school funding was larger than the additive impact of both main effects. In contrast, analyses of the HS RCT show no significant interactions between the random offer of HS and positive features of children's kindergarten and 1st grade classroom instruction and elementary school quality (Jenkins et al., 2018).

Studies using the ECLS-K, where preschool was not randomly assigned, show mixed evidence of sustained effects of preschool under conditions of later environments characterized by different metrics of quality. Magnuson et al. (2007) found a compensatory effect: Positive impacts of pre-k participation on student academic achievement in elementary school were diminished in the context of better school environments (i.e., smaller elementary school class sizes and more reading instruction), but impacts continued to persist into elementary school for students in larger classes with less reading instruction. This pattern indicates a protective effect of pre-k against the future ill effect of larger class size. Ansari and Pianta (2018) found a complementary effect of ECE participation through 5th grade for children attending moderate-quality and high-quality elementary schools, but not low-quality elementary schools. In that study, school quality was represented by an aggregate index of 20 indicators measured across four grades and averaged over time. This pattern indicates dynamic complementarity. In contrast, Bassok, Gibbs, and Latham (2018) examined elementary school achievement, instruction, teacher education, and several other factors in both ECLS-K waves and found no evidence of sustained impacts in higher-quality settings.

Experiences after the pre-k year have proven important in qualifying the otherwise dismal findings of the TNVPK RCT. Swain et al. (2015) used data from the TNVPK RCT study to assess whether being exposed to a high-rated teacher in 1st grade could help sustain pre-k effects. Swain and colleagues reported some positive interactions between assignment to TNVPK and later teacher quality, though these effects were small and not observed across all outcomes. Relatedly, Pearman et al. (2020) examined a subset of children participating in the TNVPK RCT and found evidence of a positive three-way interaction among schools regarding quality, teacher quality, and access to pre-k. In other words, they found that assignment to the TNVPK program produced positive impacts on Grade 3 achievement in both math and language arts when students were subsequently exposed to higher quality schools and teachers.

In our analyses, we shed new light on the sustaining environments hypothesis by testing moderation by a host of school characteristics that capture the quality of the elementary school where children were enrolled during 5th grade. These school characteristics are described in greater detail in our methodology section but include both measures of the general school academic environment and measures of teacher quality and funding.

Post Pre-K Economic Environments

The economic conditions of childhood have meaningful impacts on each child's development. For example, evidence from the Great Recession documents that children living in areas with high unemployment rates experienced worse mental health and higher rates of special education placement (Golberstein et al., 2019). For young children (<13 years), random assignment to a higher income neighborhood in the Moving to Opportunity study improved college attendance and earnings in the long-run (Chetty et al., 2016). In terms of academic achievement, Ananat, Gassman-Pines, and Gibson-Davis found that county-level layoffs and plant closures negatively affected children's achievement test scores (2011), and mental health in adolescence (2014).

We test whether economic conditions operate independently of pre-k or in interaction. Pearman (2020) examined whether local economic conditions moderated the effect of the TNVPK program. They found a significant compensatory interaction effect: among children living in high-poverty neighborhoods, the effect of random assignment to TNVPK was positive and significant in 3rd-grade reading achievement (about a half *SD* higher), but also found that the program had adverse effects on children living in lower-poverty neighborhoods. This same interaction effect did not hold for math achievement.

We build on this literature and capture the economic conditions experienced during children's elementary school years. Adopting the approach employed by Ananat et al. (2011), Gassman-Pines et al. (2014), we test whether county job-loss alters the impact of NC Pre-K on children's achievement. We then use three other measures of county economic conditions (% enrolled in the Supplemental Nutrition Assistance Program [SNAP], % enrolled in Medicaid, and median family income) to test for heterogeneity of the NC Pre-K effect.

Summary

The evidence for pre-k effect heterogeneity suggests that specific features of the child and their environment may explain why some early programs have more persistent positive long-term effects than other programs. The theories of complementarity, compensatory effects, and additive effects all provide plausible explanations for how pre-k experiences might interact with other environmental experiences children face before, during, and after pre-k. In the following chapter, we outline our empirical strategy using data covering the NC Pre-K Program to evaluate the three theories and provide evidence on the long-term developmental consequences of NC's investment in public pre-k.

IV. Methods

As we detailed in Chapter I, our study has two overarching goals: (1) update and extend previous analyses of the long-term effects of the NC Pre-K Program on children's key academic outcomes; (2) examine whether patterns of heterogeneity comport with theories of dynamic complementarity, compensatory effects, or additive effects. Based on previous work (Dodge et al., 2017; Ladd et al., 2014), we expected again to observe positive impacts of increases in funding for the NC Pre-K Program on child achievement measured in late elementary school. However, for the examinations of environmental heterogeneity, our hypotheses were not as solidified. As Chapter III reflects, we recognized patterns consistent with complementarity, compensatory effects, and additive effects throughout the literature, making it difficult to make firm a priori predictions about what the particular interactions examined here would produce. However, we made a priori decisions about which moderators to pursue, and as we detail below, we attempted to standardize our empirical approach across the various moderation tests as much as possible (though data restrictions required some affordances in several specific cases).

Our study builds upon the data and analytic strategy described in Dodge, Ladd, Muschkin and colleagues (2017, 2014, 2015), which estimated the effect of variation in funding for the program on age-eligible children across the state. In this study, we included an additional six birth cohorts of children who were not part of the 12-cohort sample in the earlier evaluations (i.e., those attending NC Pre-K between the 2004 and 2010 school years) and we incorporated a rich set of early childhood, elementary school, and county-level economic environmental factors derived from various sources to test for moderation. We used reading and math academic achievement scores (combined into a single variable), special education status, and retention status in 5th grade as our study's outcome measures to capture children's cumulative exposure to elementary school environments.

Participants

The sampling frame for the study included all live births in NC, as captured by individual vital statistics records from October 17, 1987 (corresponding to the first birthdate eligible for NC Pre-K in 1991–1992), to August 31, 2005 (corresponding to the last birthdate eligible for NC Pre-K in 2009–2010), from the NC Division of Public Health. These records were then linked with state educational administrative data to connect each child's early childhood environment (county of birth) with their later educational outcomes.

The process of linking individual birth records to individual public-school educational records was done securely by the North Carolina Educational Research Data Center (NCERDC) at Duke University under IRB# [2017-0495]. Of the 1,918,794 births appearing in our records during this period, 1,425,408 were matched as attending a public school in NC in any grade and year (74.29% of children born in NC during this period). Non-matches occurred because of migration out of state, private school attendance, or errors in the matching process. The matched group was slightly more likely than the unmatched group to have a mother who is single, native born, and non-White. Importantly, match rates were not systematically related to county funding allocations, alleviating concerns of biased sample selection, and birth-record variables (see list below) were included as covariates in analyses (see Ladd et al., 2014, for more details about the matching process). Among them, 1,251,295 students attended 5th grade in a NC traditional noncharter school. Students who attended a charter school were excluded from our analysis sample because their educational experiences in elementary school are anomalous to traditional public-school settings and because several measures of the school environment are not collected from these schools. Finally, we exclude students with incomplete birth record information. This results in an analysis sample of 1,248,937 students, 1,207,576 of them with valid 5th grade math and reading achievement data.

Measures

The outcomes and independent variables that are used across all analyses are presented in this subsection, with descriptive statistics shown in Table 1. The moderators are presented in the following subsection, and summarized in Table 2. We present as much demographic detail as the data files afford. The files identify a student as Hispanic even if they also identified with another race/ethnic groups (e.g., a Black Hispanic student was coded as Hispanic). The majority of students in our sample were identified as non-Hispanic White (58%), 29% non-Hispanic Black, 2% non-Hispanic Native American, 1% non-Hispanic Asian, 3% non-Hispanic multiracial, and 7% Hispanic (any race). Predictably, 50% of students were coded as female and 50% as male (no other gender group categories were available). Of all children in the sample, 8% were low birthweight, 65% came from households with parents who were married, and 24% had mothers who had less than 12 years of education. The elementary schools that children attended had a mean daily membership of 576 students, and on average more than half of the students were economically disadvantaged (54%).

TABLE 1

Descriptive Statistics of Treatment and Control Variables for Analytic Sample

	Mean	*SD*	*N*	Min	Max
Academic composite (reading and math scores, avg, std)	0.00	1.00	1,207,576	−4.31	3.40
NC Pre-K funding ($000s)	0.43	0.69	1,207,576	0.00	4.87
NC Pre-K funding ($000s; no zero values)	0.86	0.75	609,273	0.01	4.87
Smart Start program					
Smart Start funding ($000s)	1.46	1.08	1,207,576	0.00	4.83
Smart Start funding ($000s; no zero values)	1.75	0.95	1,012,294	0.00	4.83
Child characteristics					
Male	50%		1,207,576		
Non-Hispanic White	58%		1,207,576		
Non-Hispanic Black	29%		1,207,576	0%	100%
Non-Hispanic Native American	2%		1,207,576		
Non-Hispanic Asian	1%		1,207,576		100%
Non-Hispanic Mixed race	3%		1,207,576		100%
Hispanic (any race)	7%		1,207,576	0%	100%
Extremely low birthweight	0%		1,207,576	0%	100%
Very low birthweight	1%		1,207,576	0%	100%
Low birthweight	7%		1,207,576	0%	100%
Normal birthweight	82%		1,207,576	0%	100%
High birthweight	10%		1,207,576	0%	100%
Mother characteristics					
Education <12 years	23%		1,207,576	0%	100%
Married	65%		1,207,576	0%	100%
Immigrant	9%		1,207,576	0%	100%
Age at child's birth (years)	26.06	5.93	1,207,576	10.00	54.00
No father information	14%		1,207,576	0%	100%
First birth	43%		1,207,576	0%	100%
White	68%		1,207,576	0%	100%
Black	29%		1,207,576		100%
Native American	2%		1,207,576		100%
Asian	1%		1,207,576		
Other races	0%		1,207,576		
Hispanic	7%		1,207,576	0%	100%
School characteristics					
Non-Hispanic Black students (%)	28%	23%	1,206,896	0%	100%
Non-Hispanic White students (%)	55%	27%	1,206,896	0%	100%
Non-Hispanic Other race students (NA, Asian, mixed race) (%)	6%	9%	1,206,896	0%	98%
Hispanic students (any race) (%)	11%	12%	1,206,896	0%	84%
Economically disadvantaged students (%)	54%	24%	1,206,283	0%	100%
School membership	576.68	199.71	1,206,896	4.00	1,558.00
School district characteristics					
Federal per pupil expenditures ($2019)	1,094.34	410.32	1,207,574	367.61	4,116.78

(*Continued*)

TABLE 1. (*Continued*)

	Mean	*SD*	*N*	Min	Max
Local per pupil expenditures in school district ($2019)	2,281.14	748.97	1,207,574	444.07	6,558.33
State per pupil expenditures ($2019)	6,381.23	671.24	1,207,574	5,333.84	16,029.91
County characteristics					
Births to Black mothers (%)	26%	14%	1,207,576	0%	81%
Births to White mothers (%)	71%	16%	1,207,576	0%	100%
Births to mothers of other racial groups (NA, Asian, or other) (%)	3%	0%	1,207,576	0%	49%
Births to Hispanic mothers (%)	8%	11%	1,207,576	0%	100%
Births to mothers with <12 years of Ed (%)	23%	7%	1,207,576	7%	49%
Number of births (log)	7.51	1.09	1,207,576	3.09	9.50
Total population (log)	11.77	1.00	1,207,576	8.24	13.59
Median family income	682.90	132.40	1,207,576	335.01	1015.74
SNAP recipients (%)	7%	4%	1,207,576	2%	24%
Medicaid recipients (%)	15%	6%	1,207,576	3%	35%

Note. Sample sizes reflect the analytic sample for academic achievement analyses (i.e., students with valid 5th grade test scores). Special Education and grade retention analyses in Chapter III include a larger sample of children who were exempt from testing. Due to how race and ethnicity are categorized in school records, we cannot identify the race of Hispanic students. However, maternal characteristics were drawn from birth records, which included race and ethnicity as separate variables.
NA = Native American; NC Pre-K = North Carolina Pre-K; *SD* = standard deviation; SNAP = Supplemental Nutrition Assistance Program.

Child Outcomes

Data on student end-of-grade (EOG) mathematics and reading achievement tests scores (NC's statewide summative assessments for students in Grades 3–8) during 5th grade came from the NC Department of Public Instruction (NCDPI), and were obtained through the North Carolina Education Research Data Center (NCERDC) at Duke University. Test scores were standardized by subject, school year, and grade to account for changes in the test over time. The mean and *SD* of the reading and math EOG test scores are presented in their original (raw) form, for each study year, in Table A1.

Given a large bivariate correlation between the standardized mathematics and reading scores (.74), we created an *academic composite* outcome variable that is the average of the student's math and reading standardized EOG scores (coefficient $\alpha = .85$), which we then re-standardized, such that the unit for all the test score outcomes is a *SD* on the composite score. We use this achievement outcome measure in our primary moderation analyses to reduce the total number of significance tests, and because we did not have a priori hypotheses about whether moderation effects would be larger or smaller on tests of math versus reading. Because our study includes 15 environmental moderators and 10 population subgroups, decreasing the possibility of chance findings was a clear priority. However, because we understand that many readers may be interested in understanding how the results reported

TABLE 2

Moderator Variable Measurement Detail and Descriptive Statistics Across Study Period

	Level of Measurement	Timing of Measurement	Time-Varying or Invariant	Mean	*SD*
Alternative ECE services					
HS saturation (%)	County	Age 4	Varying	10.70	5.96
Early Head Start saturation (%)	County	Age 0–2 (avg)	Varying	0.32	1.32
Smart Start funding ($000s)	County	Age 0–5 (avg)	Varying	1.46	1.08
Head Start presence (0/1) at NC Pre-K start	County	2003	Invariant	0.72	0.45
Head Start saturation (%) at NC Pre-K start	County	2003	Invariant	10.29	4.95
School characteristics					
School-average achievement composite (lagged; std.)	School	Grade 5	Varying	0.02	0.99
Local PPE in 000s ($2019)	School-district	Grade 5	Varying	2.28	0.75
State PPE in 000s ($2019)	School-district	Grade 5	Varying	6.38	0.67
Federal PPE in 000s ($2019)	School-district	Grade 5	Varying	1.09	0.41
Teacher characteristics					
Student–teacher ratio in the school	School	Grade 5	Varying	15.50	3.50
National Board Certified teachers (%)	School	Grade 5	Varying	11.13	9.07
Teachers with <3 years experience (%)	School	Grade 5	Varying	21.58	10.69
Annual teacher turnover (%)	School	Grade 5	Varying	15.55	9.51
County economic factors					
Job loss (% affected)	County	Grade 5	Varying	0.87	0.84
Median family income ($00s)	County	Grade 5	Varying	676.76	128.94
SNAP recipients (%)	County	Grade 5	Varying	12.24	6.25
Medicaid recipients (%)	County	Grade 5	Varying	19.74	6.20
Observations	$n = 1{,}207{,}576$				

Note. ECE = early childhood education; NC Pre-K = North Carolina Pre-K; PPE = per-pupil expenditure; *SD* = standard deviation; SNAP = Supplemental Nutrition Assistance Program.

here extend to the individual domains of math and reading, results for each outcome component measure are shown in Tables A6–A13.

Moving beyond measures of achievement, we created dichotomous indicator variables for students' *special education placement based on exceptionality status* (excluding those academically or intellectually gifted) between their 3rd- and 5th-grade year (1 = received special education services), and for their *retention status* between their 3rd- and 5th-grade year (1 = retained between 3rd and 5th grade). Unfortunately, our data did not contain information regarding retention or special education placement prior to the 3rd-grade year.

NC Pre-K Funding

Key to our identification strategy is measuring variation in the annual funding allocations for NC Pre-K awarded to each of NC's 100 counties, derived from administrative records for the program. All funding levels were scaled to thousands of dollars and were converted to 2019 dollars. Our study's primary measure of pre-k exposure is equivalent to the level of NC Pre-K funding (per 4-year-old child) in the county in a year. The per-child dollar allocations for each of the 100 counties across each of the 18 birth cohorts included in the study were matched to specific children based on their county of residence at birth.[3] Assigning NC Pre-K funding exposure by children's birth location could be problematic if families systemically move between birth and age 4, when they become eligible for NC Pre-K. Although we did not have information on each child's pre-k location, we did have their elementary school location, which we used to calculate the number of students who moved across counties. Approximately 24% of children in our matched sample moved between birth and 5th grade. We also tested whether our main analyses were robust to the exclusion of students who moved, findings for which are presented in Chapter V.

Funding was allocated annually to each of the 100 NC counties to support classroom-based slots for eligible children (i.e., the program did not fund separate preschool programs). Funding began in pilot counties in fiscal year 2002 and varied in the number of counties and dollars per student across years (Ladd et al., 2014). In the first year, 34 of 100 counties received funds. In the second year, 57 counties were added, and in the third year all 100 counties received at least some funds.

After funding was allocated to a county entity that administered the funds, a funding slot was awarded to a pre-k classroom if the classroom met standards set by the state, designed to improve the overall quality of pre-k. As we detailed in Chapter II, these standards included staff qualifications, class size, teacher–child ratios, and North Carolina child care licensing requirements. Furthermore, through this requirement, the state's goal was to provide high-quality pre-k not only to the funded children but also to other children enrolled in the same centers. Of the pre-k classrooms enrolling at least one funded child, about a third of the children did not receive state funding. This funding also went to a variety of settings. For example, in 2004–2005, 50% of the slots were in public school settings, 31% in for-profit community child care centers, 11% in nonprofit child care centers, and about 10% were in HS programs.

The per-child NC Pre-K funding level for each county, over the full course of the study period, is displayed in Figure 3. This level is computed as the actual total number of dollars allocated to the county in a year divided by the estimated number of 4-year-old children living in that county in that fiscal year, without regard to the funding amount allocated to each funded child. This figure illustrates the within-county variation that provides the identification of the NC Pre-K funding effect. Beginning in 2002, program funding was allocated at different amounts between and within counties across years. By the 2010 program year, the state provided an average of approximately $2,030 per 4-year-old child in the county (averaged across counties), and the program enrolled 28% of all 4-year-olds in NC. The average per-child NC Pre-K investment from state allocations across all of the years when the program was operating in our study (2002 through 2010) was about $1,064 per 4-year-old child in a county. Local NC Pre-K program contractors in counties were also required to supplement state allocations with some additional funds in order to fund each individual slot fully, but we are not able to measure these local funding supplements in this study. The average within-county variation in funding once they started disbursing funds was $798 per 4-year-old. Evidence in support of the conditional

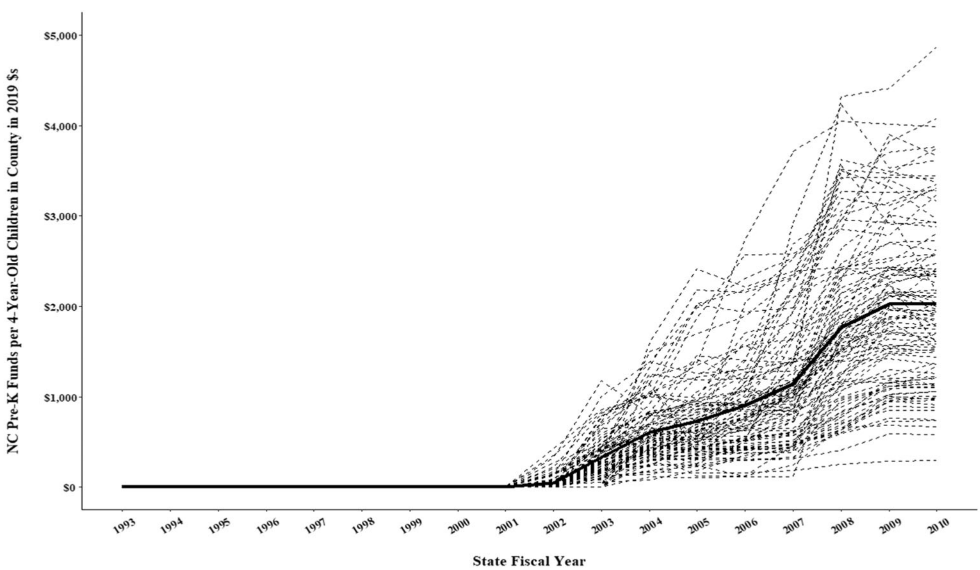


Figure 3.—North Carolina Pre-K (NC Pre-K) funds per 4-year-old child in county: 1993–2010.

exogeneity of NC Pre-K funding allocations (controlling for county population characteristics related to NC Pre-K eligibility) is presented in Chapter V.

Environmental Moderators of NC Pre-K

Our analyses are organized by the timing of each moderator relative to NC Pre-K; factors that were co-occurring with NC Pre-K during early childhood, and those occurring after NC Pre-K during elementary school. Moderators occurring after pre-k exposure are further organized into conceptual groups by school, teacher, and county economic factors. We present the description of our moderators and findings in this order, and summarize our moderator variables, their level of measurement, time-varying nature, and mean value across all study years in Table 2. Pairwise correlations for each of the study moderators is available in Table A2. We present descriptive statistics of our moderators in their original units. For the moderation analyses presented in Chapter VI, we grand-mean-centered the values for each continuously measured moderator to ease interpretation of the main effects in the presence of interactions, and converted measures in percentage units to 10 percentage point units such that regression coefficients represent more interpretable increases in moderator variables.

Early Childhood Educational Environment

During the early childhood educational period, we tested for moderation of NC Pre-K based on the access or availability to alternative forms of ECE in the community. Specifically, we operationalize access and availability as the level of funding and enrollment for several other state- and federally-funded early childhood programs in each county. Here, we briefly describe how we calculated these measures.

Smart Start funding. Described in previous evaluation reports (see Dodge et-al., 2017; Ladd et al., 2014), the Smart Start program was another state-level investment in early childhood services operating in the years of our study. Smart Start was designed to support children from birth to age 5, with funding provided for childcare vouchers and services, professional development, and health and family support services. Smart Start began in fiscal year 1994 in pilot form in 18 counties (two sites within each of North Carolina's nine Congressional districts at that time), increased to more than 50 counties by 1997 and was implemented in all 100 NC counties by the 1998–1999 school year. Beginning in 1994, the average cumulative exposure to Smart Start funding grew steadily and reached a peak of $2,383 per child (calculated as $477 per 0- to 4-year-old child per year × 5 years of exposure) in 2004, but then showed decline to a value of $1,853 per child in 2010 (the last year of program funding considered in the current study). The average cumulative exposure to Smart Start funding between state fiscal years 1994–2010 was $1,539 per child. Children could benefit from Smart Start funds for 5 years across ages 0 to 5. Thirty percent of all local Smart Start dollars were required to be spent on child-care services for qualifying low-income children.

In our analyses, we measured Smart Start "access" or availability at the county level in the same fashion as exposure to NC Pre-K, by using the county-level funding for Smart Start in a given year. However, because Smart Start services were offered from ages 0 to 4, we calculated the sum of a given child's level of exposure to funding across this 5-year period and used the child's county of birth to determine county-level funding for Smart Start. Following the example from previous reports (i.e., Dodge et al., 2017; Ladd et al., 2014; Muschkin et al., 2015), we controlled for Smart Start funding at the county level in each of our key analyses. We then also examined the interaction between Smart Start funding and NC Pre-K funding to test for moderation.

EHS and HS availability. We consider how access to the federally-funded EHS and HS programs might affect impacts of the NC Pre-K program on children's elementary school outcomes. To measure HS and EHS availability, we combined the linked birth-school-county records to county-level Head Start Program Information Report (PIR) data from the 1988–2010 school years. The PIR data are grantee-level administrative data collected annually by the Office of Head Start to describe the services, staff, children, and families served by all HS programs. Program staff at each grantee office submit data to the Office of Head Start monthly, and the PIR data represent an annual account of program information. Importantly, these data provide the zip code for the grantee's location, which allowed us to assign each grantee to a service region. On average, we observed one grantee per county, ranging from 0 to 2.

Unlike school districts with strict residential and geographic boundaries, EHS and HS programs are not restricted to serve only the children in the grantee's county. Thus, grantees may be based in one county and fund programs in multiple counties across the state. Unfortunately, the PIR data provide zip code information for the grantee only, and not for each individual center operated by that grantee. To address this issue of service location versus grantee location, we created a set of HS clusters that contain all counties where that grantee's funding is allocated. Based on a detailed review of all HS and EHS center locations in 2021 using the Office of Head Start's online directory, as well as contacting each grantee individually to confirm their service area history,[4] we determined the service region for each HS and EHS grantee (i.e., the counties where each grantee was operating HS and EHS centers). We then disaggregated the count of children enrolled in EHS (age 0, 1, and 2) and HS (age 3 and 4) across the counties within each grantee's service region based on the population of age-eligible children in each county. County population estimates for 0-4-year-old children were obtained from the North Carolina Office of State Budget and Management

(NCOSBM). The Surveillance, Epidemiology, and End Results Program (SEER) population estimates were used when NCOSBM estimates were not available. See Appendix B for additional details on our HS and EHS cluster methodology and variable information, and comparative descriptive calculations derived from the PIR and demographic data with the cluster methodology.

We used county- and year-level 4-year-old HS enrollment, divided by the total population estimate of 4-year-olds in the county that year, to calculate the proportion of children enrolled in HS in each county (i.e., HS enrollment "saturation"). For example, a county with 100 4-year-olds enrolled in HS and a population of 1,000 4-year-olds would have been assigned a HS value of .10, or 10%. As a robustness check, we also created a version of the HS exposure variable based on the average of the PIR and Head Start Cluster calculations. The HS exposure variable is divided by 10 to ease interpretation of coefficients (i.e., coefficients can be interpreted as a 10-percentage point increase in cumulative HS exposure).

To measure access to EHS, we first took the maximum value of county-level EHS enrollment of 0-, 1-, and 2-year-old children enrolled in EHS across each 3-year period in which the child would have been age-eligible to participate in EHS. We then divided this maximum amount by the total population estimate of this age group in the county during each year to calculate EHS saturation. We used the county-level maximum value of EHS enrollment by age, rather than the mean, to account for the fact that EHS programs often enroll more 2-year-olds than new infants (i.e., a more generous approach). We also created a measure of EHS saturation that uses the *mean* value of 0–3-year-old enrollment in each county, rather than the maximum value. Both of the EHS exposure variables are divided by 10 to ease interpretation of coefficients (i.e., coefficients can be interpreted as a 10-percentage point increase in cumulative EHS exposure).

We also constructed a HS measure that would allow us to test whether the presence of HS in a county at the beginning of the NC Pre-K program made the county "shovel ready" for NC Pre-K scale-up, providing a "springboard" for the NC Pre-K's program success (e.g., ECE infrastructure that facilitated scale-up). We created two additional variables to test this possibility. The first is a basic indicator variable that equals 1 if HS was present in a county in 2003, the first year of NC Pre-K's statewide availability. The second is the 2003 level of age-4 HS saturation. Although these two measures are not time-varying and would drop out of the county fixed effects regression on their own, we can recover an estimate of the interaction of each measure with NC Pre-K funding, allowing us to test for moderation.

School and Teacher Environments

We next examined whether elementary school characteristics moderate the long-term impacts of NC Pre-K funding exposure. Our measures of students' elementary schools and teachers were derived from annual administrative data from NCDPI, made available through Duke University's NCERDC. Each measure represents the elementary school conditions in a student's 5th grade year, unless otherwise noted. We defined elementary schools as traditional public schools (excluding charter schools) that included 5th grade in any primary school configuration (e.g., schools with grade ranges K-5, pre-k-6, 3–5), the year our outcomes were measured.

School average achievement. We began with a broad measure of school-level academic achievement. To create this measure, we first calculated an academic composite at the school level by averaging the standardized school-level average of Grade 5 math and reading EOG scores. This measure was calculated based on all available student test score information for a given school (regardless of whether the student was included in our analysis sample). Then, for each student, we assigned the value of the year *prior* to the student being in 5th grade (i.e., for a child attending 5th grade in year t, the moderator values are drawn for the 5th-grade scores in her school in $t - 1$). This adjustment assures that a given student's own 5th grade score (i.e., the outcome variable) would not contribute to the moderator measure (i.e., the independent variable). This measure is missing the first year a school is observed.

Federal, state, and local per-pupil expenditures (PPEs). Per-pupil funding was calculated at the school district level (for all Grades K–12, in 2019 dollars) for each school year, financed by the federal, state, and local governments. To be consistent with our NC Pre-K funding measures, we then scaled these measures to thousands of dollars. School districts and counties are often the same geographic identities in NC, with some counties containing more than one district (100 total counties and 117 school districts).

Student–teacher ratio. This variable is the total number of students in the student's elementary school (defined as the school in which the child was enrolled in 5th grade), across all grades, divided by the total number of teachers (including special education teachers), across all grades, for each school year. A higher ratio would imply larger average class size.

National Board Certified Teachers. For this measure, we calculated the proportion of teachers in the school holding a National Board Certification (NBC). This was derived by taking the total number of NBC teachers in the student's elementary school across all grades and dividing by the total number of teachers in the school for each school year. As with other percentage measures, we then converted the measure to include 10 percentage-point units to ease interpretability of coefficients. Note that the source of these data changed over time: NCDPI school report card data for 2002–2005, NCDPI administrative records for 2006–2018, and interpolated from school-level averages for 1998–2001.

Inexperienced teachers. This is the proportion of teachers in the student's elementary school who have less than 3 years of teaching experience. Similar to the NBC measure, this was calculated by dividing the total number of inexperienced teachers across all grades by the total number of teachers in the school for a given school year. We then converted the measure into 10 percentage-point units.

Teacher turnover. This measure captures the proportion of classroom teachers in the student's elementary school who left the school in a given year, calculated by dividing the total number of teachers across all grades who left the school in the past year by the total number of teachers in the school. This measure was also converted into 10 percentage-point units.

County Economic Environment

Our final set of moderation analyses examined county-level economic conditions during students' 5th grade year.

Job loss. We merged data on all business closures and layoffs, derived from the Business Closings Database, taking place within all 100 counties in North Carolina from 1997 to 2011. We then constructed a measure that captures the percent of working-age adults affected by the total yearly job loss in a given county and year due to business closures. This was calculated by taking the total yearly job loss from July to June of each year (aligned with the corresponding school year of the child's Grade 5 test) and dividing by the total number of working-age adults in that county (age 25–64). A more detailed description of these data can be found in Ananat et al. (2011; includes data through 2007 only).

We used three additional indicators of the county-level economic environment for our moderation analyses, each derived from the North Carolina Office of State and Budget Management (NCOSBM) for the calendar year of students' 5th grade year.

Median family income. Estimates of county-level median family income were obtained from Housing and Urban Development (HUD) and converted to 2019 dollars for each calendar year.

Percent of SNAP recipients. Derived from NCOSBM in a county in each year, this is the percent of the population who are receiving SNAP (formerly known as food stamps) at the county level for each calendar year.

Percent of medicaid recipients. Derived from NCOSBM in a county in each year, this is the percent of the population who are Medicaid recipients at the county level for each calendar year.

We note that the SNAP and Medicaid measures are highly correlated (see Table A2), but have some distinct features. SNAP is an entitlement program, and is not subject to state budgetary restrictions, whereas Medicaid is not an entitlement program, and eligibility and generosity is subject to state discretion. Unlike these two measures, median family income captures county economic productivity, rather than need.

Control Variables

As we describe below, all models relied on two sets of fixed effects (county and year) to isolate "exogenous" variation in NC Pre-K funding on child outcomes. We also included a set of child, family, and county characteristics in all models presented to increase precision and control for any time-varying county effects that might influence both NC Pre-K funding and child achievement, which are not controlled by county fixed effects.

Child and Family Characteristics

Derived from the matched birth and school records, we included the following child characteristics as covariates in our analyses: the infant's quarter of birth (as fixed effects), sex, birthweight (extremely low [<1000 g], very low [1000–<1500 g], low [1500–<2500 g], normal [2500–<4500 g], high [<4500 g], a single exhaustive and mutually exclusive categorical indicator for race/ethnicity (non-Hispanic Black, non-Hispanic Native American, non-Hispanic Asian, non-Hispanic Mixed race, non-Hispanic White, and Hispanic). We also controlled for mother characteristics including: marital status (married = 1, unmarried = 0), years of education, age at child's birth (years), primipara (yes = 1), race (Black, Native American, Asian, White, other race), ethnicity (Hispanic or not Hispanic), and immigration status (1 = yes). Finally, we also controlled for whether a father was present at birth (no = 1). It should be noted that for child data, information on race and ethnicity was provided by the schools in a form that coded "Hispanic" as a mutually exclusive category with the other race/ethnic groups. The maternal data were derived from the birth records, and it included ethnicity as a separate variable apart from race. Because these variables were controls in our models, we kept the coding as it was provided by our data sources. Finally, because regression analyses require that one group in a series of mutually-exclusive indicator variables serve as the omitted-reference category, we used the largest group (i.e., White) as the omitted category in our models.

Birth County Characteristics

Our analyses included a host of time-varying county characteristics matched to each child's observation by their year of birth. These measures were derived from aggregating information from all birth records in a given county and year, and from the NCOSBM: percent births to non-Hispanic Black mothers, percent births to Hispanic mothers, percent births to low-education mothers, total population (log), number of births (log), proportion of population receiving SNAP, proportion of population receiving Medicaid, and median family income (in 2019 $). Note that the latter three measures were used as moderators in the economic conditions analyses. It should also be noted that because these variables were measured at an aggregated level, we include fewer controls. Thus, for demographic indicators of population race and ethnicity, we only include controls for births to non-Hispanic Black mothers and Hispanic mothers, respectively, as those measures likely capture factors related to systemic disadvantages.

Analytic Plan to Estimate NC Pre-K Main Effects

We followed the identification strategy set out by Ladd et al. (2014), whereby we regressed 5th grade outcomes measured at the student-level on the levels of state financial allocations to the county for NC Pre-K in the fiscal year(s) when the child was age-eligible to benefit from the presence of the program. This natural experiment exploits two different sources of variation. First, we leveraged between-county variation in the timing of funding receipt, as not all counties received the same amount of funding for NC Pre-K in every year (shown in Figure 3 for all program years included in our study). Second, we leveraged within-county variation in NC Pre-K funding over time, as counties received different NC Pre-K funding allocations across the years considered here. Therefore, children from the same birth cohort who were born in different counties would have different levels of "exposure" to NC Pre-K according to our funding variable and modeling approach. Moreover, children from different birth cohorts who were born in the same county would also have been exposed to different levels of NC Pre-K funding. This allowed us to compare children across years, before and after funding changes, and living in the same county, to assess the impact of funding.

Our first set of analyses examined the main effect of NC Pre-K for our updated analytic sample using a basic OLS regression model specification that included two-way fixed effects for county and program year:

$$E_{icty} = \delta NCPK_{cy} + \boldsymbol{X}_{ct}\alpha + \boldsymbol{Z}_{i}\beta + \theta_{y} + \pi_{c} + \varepsilon_{icty}, \tag{1}$$

where E_{icty} denotes the educational outcome (i.e., achievement, special education, or grade retention) for child i born in county c in year t who was eligible to participate in NC Pre-K at age-four in year y. $NCPK_{cy}$ is a continuous variable representing the NC Pre-K funding allocations for county c in program year y. X_{ct} represents the vector of time-varying county covariates listed above, including Smart Start funding allocations, and Z_i represents a vector containing the full list of individual child-level covariates (e.g., the education level of the child's mother, birthweight of the child). To capture potential developmental differences based on age between children in a cohort (i.e., at the time of pre-k enrollment, children can vary in age by a year based on age-eligibility cutoff requirements), we also included birth-quarter fixed effects for each child in Z_i. Importantly, the program year fixed effects, θ_y, defined for each student by the academic year for which they were age-eligible for NC Pre-K, capture any unobserved shock common to all children during the same program year.

Previous work on NC Pre-K using this design (e.g., Dodge et al., 2017; Ladd et al., 2014) relied on birth year fixed effects, which would capture differences between calendar years in factors that might influence student achievement. Because our work was solely focused on estimating NC Pre-K effects (previous work also prioritized modeling impacts for the Smart Start program alongside NC Pre-K effects), we updated our approach to use program year fixed effects, which groups students based on the *school year* for which their funding level is allocated. Thus, our set of fixed effects should more accurately capture potential factors in a program year that could affect a cohort of students who experienced the same level of funding for NC Pre-K. Finally, the county fixed effects, π_c, control for observable and unobservable time-invariant characteristics of the county, which helps to address the potential for endogeneity of NC Pre-K funding allocations during policy adoption, and restricts comparisons of a child's outcomes to other children within the same county over time. This strategy creates a better “apples to apples” comparison than standard linear regression because it compares children who lived in the same communities but were differentially exposed to NC Pre-K based on their timing of birth. In turn, the county and year fixed effects together allow us to estimate a version of a Difference-in-Difference model, comparing counties to themselves over time, before and after major policy changes, and then averaging those within-county differences across the state. We estimated all linear regression models in Stata 17.0 using the *reghdfe* package (Correia, 2017). The dichotomous outcomes of special education and retention status were estimated with logistic regression, with results converted into odds ratios. Robust standard errors were adjusted for clustering by county.

An important assumption underlying our empirical strategy is the exogeneity of the funding level of the programs in a particular county year. When all the covariates are included in the model, our approach estimates the association between the funding level to which a child was exposed and that child's outcomes. This estimate is based on integrated comparisons between children growing up in the same NC county but in different years, whose exposure to pre-k funding differed, and between children born in the same year but living in different counties, again whose exposure to pre-k funding differed. When making this estimate, we hold constant all of the child-level and county-level covariates listed above. Thus, our modeling approach assumes that conditional on the set of county and program year fixed effects and time-varying county characteristics, the level of funding for NC Pre-K for which a given child was exposed was “as good as random.” Of course, this is a strong assumption, and rests jointly on our understanding of the way the NC Pre-K program was expanded by the state (described in Chapter II) and on our examinations of the data in hand.

Unfortunately, no set of empirical tests can fully rule out the possibility that unobserved confounds could exist that cause both year-to-year changes in county funding and student achievement. However, such confounds would need to be unrelated to the set of county-level covariates included in our model, and they would need to be changing within a county over time to not be negated by the set of county fixed effects. In previous work, Ladd et al. (2014) examined such possibilities by including a set of sensitivity tests that were designed to test for the plausibility of potential sources of bias, and they concluded that conditional on the set of county and year fixed effects, the funding differences appeared to be exogenous. Shown in Chapter V, we set forth a similar set of checks, probing sensitivity to different sets of control variables and modeling assumptions, including sets of county-specific time trends.

Finally, we created four sets of population characteristics by which to estimate our subgroup analyses to test for heterogeneity of the NC Pre-K treatment effect (Chapter V): *Race and ethnicity* (Black, Hispanic, White, and other racially identified students); *sex* (males, females); *birthweight* (low birthweight, regular/high birthweight); and *mother's education* (less than high school degree, high school degree or higher). Note that due to the small sample sizes for several of the race/ethnicity groups, we collapsed children who were identified as Asian, Native American, and Mixed race into a single group. We estimated Equation (1) separately for each of the subgroups listed and tested for statistical differences in the coefficient on NC Pre-K across groups by estimating models that interacted NC Pre-K and the subgroup characteristics (Bloom & Michalopoulos, 2013).

Analytic Plan for Key Tests of Environmental Heterogeneity

We tested for effect moderation of NC Pre-K by adding interaction terms to Equation (1) presented in Chapter V:

$$\begin{aligned} E_{icty} = {} & \delta NCPK_{cy} + \gamma MOD_{cy} + \lambda NCPK_{cy} \times MOD_{cy} + \boldsymbol{X}_{ct}\alpha \\ & + \boldsymbol{Z_i}\beta + \theta_y + \pi_c + \varepsilon_{icty}, \end{aligned} \tag{2}$$

where MOD_{cy} represents the main effect for a given moderator of interest that is centered at its mean value. The parameter of interest for these models, λ, captures the interaction between the moderator and NC Pre-K within a given county and year to determine whether the moderator affected the impact of NC Pre-K funding on a given outcome. All other terms are the same as defined above. We estimated a version of Equation (2) for each moderator in our study. Table 2 shows all of our study moderators, the level (e.g., school, county) and timing (e.g., age 4, 5th grade) at which they are measured, and the way each moderator was measured over time (i.e., time-varying or time-invariant). For each conceptually-linked set of moderator variables (e.g., school characteristics, economic conditions, etc.), we also estimated an “omnibus” moderation specification. This omnibus test included each of the conceptual groups' moderators and interactions with NC Pre-K in the model simultaneously to examine the robustness of the different measures when modeled together, and we used this specification to test for the joint significance of the conceptual grouping of the moderators.

Our final set of moderation analyses combines the exploration of subgroup and environmental heterogeneity. Here, we further probe the environmental moderators that we find significantly relate to achievement by estimating each of these moderation analyses separately for each population subgroup examined in Chapter V (e.g., race and ethnicity, mother's education) to more comprehensively explore the ecological relations between family and child characteristics and environmental experiences. We conducted these intersectional moderation analyses only for moderators that produced a statistically significant interaction term with NC Pre-K funding to reduce the number of statistical tests and thus the possibility of spurious or chance associations.

Understanding NC Pre-K's Effects for Those who Participated in the Program: An Instrumental Variables Approach

From the policymaker's perspective, analyses of the effect of NC Pre-K funding tell us how allocating resources to communities affects academic achievement for the *entire* population of age-eligible children—regardless of whether a particular child received a funded slot for NC Pre-K. This estimate is undoubtedly important and represents our preferred approach, but it does not offer an estimate of the effects on those children who received a funded slot for pre-kindergarten.

To address these concerns, we conducted supplemental analyses using individual-level NC Pre-K enrollment status for all children in our sample. Including these data requires caution; although the two-way fixed effects model should produce plausibly exogenous variation in NC Pre-K funding at the county level, the choice to enroll an individual child in NC Pre-K is not. The coefficient estimates from a regression of student achievement on enrollment status incorporates both the causal effect of NC Pre-K as well as some amount of unobserved bias from selecting to participate. If the factors that lead a family to choose pre-k enrollment also cause later achievement, then the coefficient on NC Pre-K enrollment in this model will likely be biased.

To move closer to a more accurate estimate of the impact on children who received an NC Pre-K funded slot, we applied an econometric tool, Instrumental Variables (IV) analysis, that allowed us to bring together our exogenous "policy" measure, county funding, with our endogenous "treatment" measure, enrollment, to estimate the impact of enrollment in NC Pre-K. The intuition of IV analysis is that it takes variation in the troublesome endogenous measure and matches it with variation in an instrument, and then uses only this exogenous variation in the instrument to estimate a treatment effect (Angrist & Pischke, 2008; Winship & Morgan, 1999). This methodology has been used in child development research that is applied to difficult selection bias issues such as how child care quality (Auger et al., 2014), neighborhood conditions (Foster & McLanahan, 1996), mathematics learning (Watts et al., 2018), and maternal education (Gennetian et al., 2008) affect children's development.

A variable must meet two conditions to work as a valid instrument: (1) it must be a strong predictor of the endogenous regressor of interest (inclusion); and (2) it can *only* be related to the final outcome of interest through its correlation with the endogenous variable (exclusion; Lee, 2005).[5] Our instrument is county NC Pre-K funding, and the endogenous regressor is NC Pre-K enrollment. Funding program slots in a community make possible one's enrollment in the program, and therefore we predicted a strong relation between funding and enrollment, meeting the inclusion assumption. In turn, we also assumed for this analysis that funding for pre-k slots would improve a student's outcome only through their enrollment and participation in the program, theoretically meeting the exclusion assumption.

However, the exclusion restriction, also known as the "only through" assumption, is difficult to assess because it assumes there are no alternative pathways between the instrument and the outcome. It is thus untestable, and researchers must rely on theory and descriptive calculations for support. We acknowledge that providing funding for a community's ECE could influence outcomes through other means besides enrollment, such as childcare center quality improvements or peer effects, if funding altered the preschool environment more generally. In fact, a prior publication (i.e., Ladd et al., 2014) asserts that the effect of NC Pre-K funding on children's outcomes in Grade 3 was so large, and so dispersed (that is, there was a significant effect on middle-income children who were not eligible for funded slots), that it must include spillover effects on peers who had not received funded pre-k slots. These possible pathways, and both the methodological difficulties and data requirements to test them, therefore lead us to a cautious interpretation of IV estimates in our study context. We view these analyses not as explicitly causal, but as a first estimate of the possible relation between NC Pre-K enrollment and student achievement.

Also note that IV analyses require at least one instrument for every endogenous variable in the model. If we were to conduct tests of moderation using enrollment as the treatment, this would involve the interaction of two endogenous variables (enrollment and moderator). Because we did not have an additional set of instruments for our moderators (it is often difficult to find one valid instrument), we could not carry forward this methodology for our other heterogeneity analyses.

Overall, we view these estimates as informative, compliance-adjusted estimates of NC Pre-K funding, providing a sense for how actual enrollment in NC Pre-K affected Grade 5 achievement that supplement our primary estimates of NC Pre-K funding main effects, and moderation of these effects by environmental factors. As we note in the following chapters, we do not believe these estimates should be interpreted as explicitly causal due to the potential problems with the "exclusion restriction" assumptions in this setting.

The statistical method for implementing IV is called two-stage least squares (2SLS), where the analysis proceeds in two steps. Step one, known as the "first stage," regresses enrollment on all study covariates including county-level funding allocations—the instrument. This produces a linear probability model, whereby the regression coefficient represents percentage point changes in enrollment as a result of thousand-dollar increases in NC Pre-K funding allocations. In this step, the methodological requirement is that the F-statistic from the significance test of the excluded instrument(s) should be above 10 (Angrist & Pischke, 2008). The second step regresses the outcome, student achievement, on the fitted values of enrollment from the first stage equation, and all study covariates excluding the instrument. This setup effectively isolates the changes in NC Pre-K enrollment that were shifted as a result of changes in county NC Pre-K funding (the first stage model), and uses only this source of discrete changes in enrollment to estimate the impact of enrollment on student achievement. We estimated these models using the *ivreghdfe* package in Stata 17.0 (Baum et al., 2010; Correia, 2017).

V. NC Pre-K Main Effects and Subgroup Results

In this chapter, we present newly estimated main effects of NC Pre-K funding on children's achievement at 5th grade before turning to models that examined heterogeneity based on key child characteristics. Described in Chapter IV, these estimates were generated from models that aggregated math and reading scores into an achievement composite, which was standardized for each cohort included in our sample, and for special education and retention status between 3rd and 5th grade.

We report results from two different approaches to estimation. In our preferred approach, as with previous work on NC Pre-K (e.g., Dodge et al., 2017), models estimated the effect of additional funding (now scaled in 1000-dollar units) for NC Pre-K for each child in a given county in a given year, regardless of the child's enrollment status in the program. The statewide average allocation per 4-year-old child was approximately $2030 in 2010. Local organizations provided supplemental funding that is not measured here. Statewide enrollment ranged from 1% of the 4-year-old population at the start of the program to 28% by the 2010 program year. We also report results from a second approach, which estimates the effect of receiving a state-funded pre-k slot on child outcomes, using an IV analysis.

Effects of State Funding on Outcomes of Age-Eligible Children

Main Effects

Table 3 presents our key results from the two-way fixed effects model, with achievement composite results shown in panel A, special education in panel B, and retention in panel C. Column 1 presents our preferred specification for the full sample of children included in our analysis (n = 1,248,937 students for the retention and special education outcomes, and 1,207,576 students with test score information). We find that an additional $1,000 of county-level NC Pre-K funding was associated with a significant increase in 5th grade achievement of approximately 0.03 *SD*. To contextualize this estimate, this would imply that an investment of $2,030 (the average per-4-year-old investment during in the 2010 program year) would lead to a 0.06 *SD* increase in student achievement at 5th grade. Note that this estimated effect, while statistically significant, is smaller than those found in previous NC Pre-K evaluations that did not include six later cohorts (Dodge et al., 2017; Ladd et al., 2014). Additionally, the results in panels B and C (displayed as odds ratios) show no association between NC Pre-K funding

TABLE 3

Main Effects of NC Pre-K Funding for the Full Sample and by Population Subgroups

	Baseline Model (1)	Race/Ethnicity: Black (2)	Race/Ethnicity: Hispanic (3)	Race/Ethnicity: White (4)	Race/Ethnicity: Other Races (5)	Gender: Male (6)	Gender: Female (7)	Birthweight: Low (8)	Birthweight: Normal or High (9)	Mother's Education: Less Than High School (10)	Mother's Education: High School or More (11)
A. Academic achievement											
NC Pre-K funding ($000s)	0.030**	0.033*	0.044**	0.018^{+}	0.012	0.029*	0.032**	0.025	0.031**	0.046***	0.021^{+}
	(0.011)	(0.015)	(0.013)	(0.009)	(0.019)	(0.011)	(0.011)	(0.017)	(0.010)	(0.013)	(0.011)
Observations	1,207,576	349,373	82,527	706,617	69,058	608,273	599,303	97,633	1,109,742	283,408	924,168
B. Special education (logistic regression, odds-ratios)											
NC Pre-K funding ($000s)	1.015	0.987	1.024	1.034	1.021	1.022	1.002	0.996	1.017	1.030	1.009
	(0.019)	(0.023)	(0.035)	(0.021)	(0.022)	(0.017)	(0.024)	(0.032)	(0.018)	(0.022)	(0.021)
Observations	1,248,937	368,087	85,182	724,170	71,492	635,146	613,791	104,753	1,143,956	299,788	949,149
C. Grade retention (logistic regression, odds-ratios)											
NC Pre-K funding ($000s)	0.996	1.007	1.047	0.968	0.989	0.994	1.000	0.954	1.002	1.018	0.986
	(0.055)	(0.067)	(0.094)	(0.058)	(0.071)	(0.058)	(0.059)	(0.078)	(0.055)	(0.064)	(0.057)
Observations	1,248,937	368,073	85,015	724,170	71,261	635,146	613,791	104,753	1,143,956	299,788	949,149

Note. Standard errors are given in parentheses. Standard errors clustered at the county level. All models include Smart Start funding (in $000s), child and family covariates, county covariates, and county, program year, and quarter of birth fixed effects. Child and family covariates include: child's sex (male), child's race and ethnicity (Black, Native American, Asian, Hispanic, Mixed race), child's birth weight (extremely low, very low, low, high), mother's education (less than high school), mother's marital status (married), mother's immigration status, mother's age (in years), whether a father was present at birth, first born status, mother's race and ethnicity (Black, Native American, Asian, Hispanic, other race). County covariates include: percent of births to Black mothers, percent of births to Hispanic mothers, percent of births to low-education mothers, number of births (log), total population (log), median family income, percent of population receiving food stamps, percent of population enrolled in Medicaid. Models run for special education placement and grade retention used logistic regression, and sample sizes were slightly larger when compared to the achievement models since the achievement models necessitated a nonmissing test score. Note that the "Other races" subgroup shown in column 5 combines children identified as either Asian, Native American or Mixed race due to small sample sizes for each respective group.

NC Pre-K = North Carolina Pre-K.

$^{+}p < .10$.

$^{*}p < .05$.

$^{**}p < .01$.

$^{***}p < .001$.

and special education and retention status between 3rd and 5th grade. The differences in results compared with those in the prior studies, which found larger achievement effects and significant effects on retention and special education placement, are due to two factors: (1) the updated model specification, which included new sets of fixed effects capturing the program year (i.e., school year) and quarter of birth (see Chapter IV); (2) our new sample that includes six additional cohorts of students, capturing state funding variation across a more recent time horizon (two-way fixed effects analyses are sensitive to panel length [Cunningham, 2021]). Within each of the six more recent cohorts, state funding was allocated to every county, and per-child funding levels across counties varied less than in previous cohorts.

Table 4 provides analyses for reading and math test scores separately. Column 1 shows that an additional $1,000 of county-level NC Pre-K funding was associated with significant increases of approximately 0.03 *SD* in 5th-grade reading scores and 0.02 *SD* in math scores. Given the similarity of the effects for these two domains of achievement and their high correlation (.74), we proceed with our other key analyses by presenting effects only on the composite score of math and reading in the main text (with supplemental results for math and reading separately shown in the Appendix).

Sensitivity Tests

As we noted above, the key assumption of our natural experimental approach is that the variation in NC Pre-K funding to counties over time is not systematically related to unobserved time-varying characteristics of counties that may also relate to children's development (time invariant characteristics are controlled with the county fixed effects). We test the robustness of our achievement composite models to this assumption with several additional specifications, shown in Table 5. In keeping with the modeling approach adopted by previous NC Pre-K studies (e.g., Dodge et al., 2017; Ladd et al., 2014), we control for Smart Start funding in each of our main models. In the first column of Table 5, we change our preferred specification by removing the Smart Start control, which produced virtually no difference in our estimate of NC Pre-K funding. This is notable, because as we detailed above, NC Pre-K funding was often allocated directly to Smart Start partnerships. Thus, the time-varying level of investment by the state in each county for Smart Start likely constitutes one of the most important potential confounds for our models. The fact that estimates do not change when Smart Start is included or removed as a control suggests that the modeling approach is likely eliminating key sources of confounding variation.

In column 2, we remove the child and family controls from the model, also producing no difference in our main effect estimate. This suggests that key child and family characteristics, which are likely related to selection into pre-k, do not appear to affect model estimates once the fixed effects and county-level controls are included. The most substantial difference occurs

TABLE 4

EFFECTS OF NC PRE-K FUNDING ON READING AND MATH ACHIEVEMENT FOR THE FULL SAMPLE AND BY POPULATION SUBGROUPS

		Race/Ethnicity				Gender		Birthweight		Mother's Education	
	Baseline Model (1)	Black (2)	Hispanic (3)	White (4)	Other Races (5)	Male (6)	Female (7)	Low (8)	Normal or High (9)	Less Than High School (10)	High School or More (11)
A. Reading											
NC Pre-K funding ($000s)	0.033***	0.031*	0.040**	0.025**	0.014	0.036***	0.031**	0.026	0.034***	0.050***	0.025**
	(0.009)	(0.014)	(0.012)	(0.008)	(0.018)	(0.010)	(0.010)	(0.017)	(0.009)	(0.012)	(0.009)
Observations	1,200,135	346,630	81,686	703,248	68,570	602,810	597,325	96,856	1,103,081	280,345	919,790
B. Math											
NC Pre-K funding ($000s)	0.024*	0.031+	0.042**	0.008	0.010	0.020	0.029*	0.022	0.024*	0.036**	0.016
	(0.012)	(0.016)	(0.015)	(0.010)	(0.020)	(0.012)	(0.012)	(0.017)	(0.011)	(0.013)	(0.012)
Observations	1,205,965	348,703	82,455	705,842	68,964	607,396	598,569	97,384	1,108,381	282,855	923,110

Note. Standard errors are given in parentheses. Standard errors clustered at the county level. All models include Smart Start funding (in $000s), child and family covariates, county covariates, and county, program year, and quarter fixed effects. See Table 3 for details. Note that the "Other races" subgroup shown in column 5 combines children identified as either Asian, Native American or Mixed race due to small sample sizes for each respective group.
NC Pre-K = North Carolina Pre-K.
$^{+}p < .10$.
$^{*}p < .05$.
$^{**}p < .01$.
$^{***}p < .001$.

TABLE 5
Sensitivity of Effects of NC Pre-K Funding on Academic Achievement

	Dependent Variable: Academic Composite				
	(1)	(2)	(3)	(4)	(5)
NC Pre-K funding ($000s)	0.030**	0.033**	0.011	0.026	0.031**
	(0.011)	(.010)	(0.011)	(0.017)	(0.011)
Child and family controls	X		X		X
County controls	X	X			X
Smart Start funding					X
Time trends				X	
HS and EHS saturation					X
Fixed-effects	X	X	X	X	X
Observations	1,207,576	1,207,576	1,207,576	1,207,576	1,207,576

Note. Standard errors are given in parentheses. Standard errors clustered at the county level. All models include county, program year, and quarter fixed effects. See Table 3 for a list of covariates.
EHS = Early Head Start; HS = Head Start; NC Pre-K = North Carolina Pre-K.
$^{+}p < .10$.
$^{*}p < .05$.
$^{**}p < .01$.
$^{***}p < .001$.

when we remove the time-varying county control variables (e.g., % county population receiving Medicaid) in column 3, where the NC Pre-K coefficient falls to .011 and is no longer significant. Column 4 includes county trends (i.e., county × year interactions) as an alternative way of accounting for time-varying county characteristics, where the NC Pre-K coefficient decreases slightly to .026 from our preferred specification, but loses statistical significance due to the increased standard error. Combined, models 3 and 4 demonstrate that controlling for time-varying characteristics of counties (i.e., county demographic compositions) are central to meeting the assumptions of the DD design in our study's context. This is not only well-established in the methodological literature, but also aligns with the targeted eligibility of the NC Pre-K program, whereby counties with greater economic disadvantage (ED)—and thus a larger proportion of eligible children—had differential funding allocations. The fact that the coefficient on NC Pre-K funding drops when the time varying county controls are removed from the model (i.e., Column 3) likely reflects the possibility that funding flowed to more *disadvantaged* areas, suggesting that variation in funding for the NC Pre-K program in our model could carry *downward* unobserved bias due to the targeted nature of the program.

The targeting of funding could raise further concerns if it indicated that such counties also had greater access to ECE more broadly; many ECE programs are similarly targeted. We address this possibility in column 5 by including county-level saturation of HS and EHS, as well as Smart Start funding. Here, the NC Pre-K main effect remains a significant 0.03 *SD*. This suggests that once time-invariant and time-varying county controls are included, NC Pre-K funding appears to have little meaningful correlation with other key early childhood investments. All told, these tests provide evidence that counties' NC Pre-K funding allocations appear to be conditionally exogenous. Although we find these results encouraging, we note in the Discussion (Chapter VII) that caution should be applied when making causal inferences from our results given that we still depend on measured control variables to eliminate bias.

Finally, for our achievement measure, we also tested the possibility that the relation between NC Pre-K funding and achievement was nonlinear by adding a quadratic term for funding to the primary specification (not shown). The coefficient on the quadratic term in this model was .00 and not significant ($p = .814$), suggesting that the relation between NC Pre-K funding and achievement was linear for the range of funding observed in our data set. That is, the effect of funding did not asymptote within the range observed.

Heterogeneity Based on Child and Family Characteristics

Moving to our first look at heterogeneity, Columns 2 through 11 of Table 3 then replicate our main model with the sample restricted to each of 10 key demographic subgroups. A scan across the 10 subgroup models shows that NC Pre-K funding had a positive coefficient for all 10 groups. The coefficients were statistically significant for six groups, including normal birthweight children, children whose mothers did not have a high school diploma, Black children, Hispanic children, males, and females. The effects for the other four groups (i.e., children in the race/ethnicity "other" category, White children, low birthweight children, and children's whose mothers completed high school) were positive, but not statistically significant.

To test if these apparent group differences were statistically significant, we tested interactions with each key demographic characteristic (see Table 6). For race and ethnicity, we focused on the two groups that showed statistically significant subgroup effects (i.e., Black and Hispanic), and the other demographic characteristics included binary splits (i.e., gender, birthweight, and mother's education). The effect of pre-k was significantly larger for low birthweight children,[6] for children of mothers without a high school diploma, for Black and Hispanic children (compared with all other racial groups, respectively), and for females.

For special education and retention, there were no significant relations with NC Pre-K funding for any of the demographic subgroups, aligning with the null findings from the full sample analyses in column 1. We also estimated linear probability models for the full sample and each of the subgroups, regressing special education and retention separately on the full set of covariates, and found the same pattern of results. Because we found almost uniformly null results for special education and grade retention, all following analyses focus solely on the achievement composite variable.

TABLE 6
Subgroup Models With Interaction Terms to Test for Statistical Significance Across Groups

	Baseline Model (1)	Black (2)	Hispanic (3)	Male (4)	Birthweight (5)	Mother Education (6)	All Mods. (7)
Interactions							
Child is Black × NC Pre-K funding		0.031**					0.040***
		(0.009)					(0.011)
Child is Hispanic × NC Pre-K funding			0.077***				0.075***
			(0.015)				(0.017)
Male × NC Pre-K funding				−0.008*			−0.008*
				(0.003)			(0.003)
Child is low birth weight × NC Pre-K funding					0.010*		0.005
					(0.005)		(0.005)
Mother's education < 12 years × NC Pre-K funding						0.047***	0.033***
						(0.006)	(0.004)
Main effects							
NC Pre-K funding ($000s)	0.030**	0.030**	0.027*	0.030**	0.030**	0.028*	0.025*
	(0.011)	(0.010)	(0.011)	(0.011)	(0.011)	(0.011)	(0.010)
Child is Black	−0.399***	−0.414***	−0.398***	−0.399***	−0.401***	−0.400***	−0.419***
	(0.012)	(0.013)	(0.012)	(0.012)	(0.012)	(0.012)	(0.014)
Child is Hispanic	−0.296***	−0.292***	−0.356***	−0.296***	−0.297***	−0.301***	−0.353***
	(0.020)	(0.020)	(0.025)	(0.020)	(0.021)	(0.020)	(0.025)
Male	−0.101***	−0.101***	−0.101***	−0.097***	−0.097***	−0.101***	−0.094***
	(0.003)	(0.003)	(0.003)	(0.004)	(0.003)	(0.003)	(0.004)
Mother's education < 12 years	−0.343***	−0.343***	−0.343***	−0.343***	−0.344***	−0.364***	−0.359***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.007)	(0.007)
Child is low birthweight					−0.156***		−0.154***
					(0.004)		(0.004)

(Continued)

TABLE 6. (*Continued*)

	Baseline Model (1)	Black (2)	Hispanic (3)	Male (4)	Birthweight (5)	Mother Education (6)	All Mods. (7)
Child is extremely low birthweight	−0.416***	−0.417***	−0.416***	−0.416***		−0.416***	
	(0.012)	(0.012)	(0.012)	(0.012)		(0.012)	
Child is very low birthweight	−0.200***	−0.200***	−0.200***	−0.200***		−0.200***	
	(0.010)	(0.010)	(0.009)	(0.010)		(0.010)	
Child is low birthweight	−0.124***	−0.124***	−0.123***	−0.124***		−0.124***	
	(0.004)	(0.004)	(0.004)	(0.004)		(0.004)	
Child is high birthweight	0.068***	0.068***	0.068***	0.068***		0.068***	
	(0.003)	(0.003)	(0.003)	(0.003)		(0.003)	
Observations	1,207,576	1,207,576	1,207,576	1,207,576	1,207,576	1,207,576	1,207,576
p Value (interactions = 0)							.000

Note. Standard errors are given in parentheses. Standard errors clustered at the county level. All models include Smart Start funding (in $000s), child and family covariates, county covariates, and county, program year, and quarter fixed effects. See Table 3 for details.
NC Pre-K = North Carolina Pre-K.
$^{+}p < .10$.
$^{*}p < .05$.
$^{**}p < .01$.
$^{***}p < .001$.

In Table 7, we present additional heterogeneity tests that examine, on an intersectional level, subgroup differences in achievement, analyzing whether children who identify as members of multiple subgroups have varied impacts of NC Pre-K funding on academic achievement. In Column 1, we present the results of an "omnibus" test that examined whether interactions with race, sex, birthweight, and mother's education status jointly contribute to the model. Indeed, we found that the addition of these interaction terms significantly contributed to the baseline model ($p < .001$), and that each of the subgroup interactions with NC Pre-K were significant (except low birthweight). In Columns 2 through 5, we focus on specific subgroups, but add interaction tests between NC Pre-K funding and sex, birthweight, and mother's education status, respectively. Generally, we found that across these various combinations of groups, effects were still larger for Black and Hispanic children, and children from mothers with less education. For White, and the "other race" groups, effects appear smaller for boys. For Hispanic, White and children identified in the "other race" group, we found that effects were largest for those whose mothers had less than a high school education. When splitting the sample on mother's education level and on birthweight, we still found that effects were larger for Black and Hispanic students across both groups on each variable (i.e., Black and Hispanic children benefitted more regardless of mother's education level or of birthweight). The largest positive interaction effect detected was for low-birthweight children who identified as Hispanic.

Summary of Main Effects

We found a significant positive effect of NC Pre-K funding on our composite measure of achievement, which also held for reading and math scores separately. We found no effect of funding on special education placement and grade retention. We found the positive effects of increases in funding on the achievement composite variable were shared by most of the key demographic groups considered, with the largest impacts for groups that often encounter more environmental adversity (i.e., Hispanic students; students whose mothers had lower levels of education). The significantly larger effects on Black and Hispanic students than White students suggest that NC Pre-K funding reduces achievement gaps across these groups.

IV Findings

Predicting NC Pre-K Enrollment (First Stage)

As noted above, we proceeded with IV analyses only for the achievement composite variable, as special education status and grade retention produced null results in our key main effect analyses. The "first stage" model results are shown in Table 8. Each model includes all the covariates in our primary

TABLE 7

Effects of NC Pre-K Funding on Academic Achievement by Intersectional Subgroups

	Model With All Chapter Moderators	Child's Race/Ethnicity				Mother's Education		Birthweight	
		Black	Hispanic	White	Other races	Less Than High School	High School or More	Low	Normal or High
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Interactions									
Child is Black × NC Pre-K funding	0.040***					0.023**	0.050***	0.039**	0.039***
	(0.011)					(0.008)	(0.011)	(0.014)	(0.011)
Child is Hispanic × NC Pre-K funding	0.075***					0.057***	0.084***	0.090***	0.073***
	(0.017)					(0.011)	(0.015)	(0.025)	(0.016)
Male × NC Pre-K funding	−0.008*	−0.000	−0.012	−0.013***	−0.018^{+}	0.006	−0.013***	−0.007	−0.008*
	(0.003)	(0.005)	(0.008)	(0.003)	(0.010)	(0.005)	(0.003)	(0.007)	(0.003)
Child is low birth weight × NC Pre-K funding	0.005	0.006	0.007	0.008	−0.013	−0.012^{+}	0.012*		
	(0.005)	(0.006)	(0.017)	(0.007)	(0.010)	(0.006)	(0.005)		
Mother's education < 12 years (1 = Yes) × NC Pre-K funding	0.033***	0.002	0.033*	0.036***	0.020*			0.004	0.036***
	(0.004)	(0.006)	(0.017)	(0.006)	(0.010)			(0.008)	(0.004)
Main effects									
NC Pre-K funding ($000s)	0.025*	0.032*	0.029^{+}	0.020*	0.011	0.035**	0.022*	0.020	0.025*
	(0.010)	(0.015)	(0.017)	(0.009)	(0.019)	(0.012)	(0.010)	(0.016)	(0.010)
Child is Black	−0.419***					−0.259***	−0.458***	−0.409***	−0.419***
	(0.014)					(0.012)	(0.011)	(0.035)	(0.013)
Child is Hispanic	−0.353***					−0.213***	−0.390***	−0.329***	−0.353***
	(0.025)					(0.020)	(0.023)	(0.052)	(0.025)
Male	−0.094***	−0.173***	−0.050***	−0.059***	−0.091***	−0.102***	−0.092***	−0.119***	−0.092***
	(0.004)	(0.005)	(0.008)	(0.003)	(0.013)	(0.006)	(0.004)	(0.008)	(0.004)

(*Continued*)

TABLE 7. (*Continued*)

	Model With All Chapter Moderators (1)	Child's Race/Ethnicity Black (2)	Hispanic (3)	White (4)	Other races (5)	Mother's Education Less Than High School (6)	High School or More (7)	Birthweight Low (8)	Normal or High (9)
Child is low birthweight	−0.154***	−0.121***	−0.139***	−0.176***	−0.152***	−0.098***	−0.169***		
	(0.004)	(0.005)	(0.018)	(0.005)	(0.018)	(0.006)	(0.004)		
Mother's education < 12 years (1 = Yes)	−0.359***	−0.261***	−0.296***	−0.431***	−0.337***			−0.320***	−0.362***
	(0.007)	(0.009)	(0.020)	(0.007)	(0.020)			(0.010)	(0.006)
Observations	1,207,576	349,373	82,527	706,617	69,058	283,408	924,168	97,633	1,109,943
p Value (interactions = 0)	.000	.794	.139	.000	.018	.000	.000	.001	.000

Note. Standard errors are given in parentheses. Standard errors clustered at the county level. All models include Smart Start funding (in $000s), child and family covariates, county covariates, and county, program year, and quarter of birth fixed effects. See Table 3 for details. Note that the "Other races" subgroup shown in column 5 combines children identified as either Asian, Native American or Mixed race due to small sample sizes for each respective group.
NC Pre-K = North Carolina Pre-K.
$^{+}p < .10$.
$^{*}p < .05$.
$^{**}p < .01$.
$^{***}p < .001$.

TABLE 8

First Stage Results From Instrumental Variables Analyses Predicting NC Pre-K Enrollment

		Child's Race/Ethnicity				Gender		Birthweight		Mother Education	
	Full Sample (1)	Black (2)	Hispanic (3)	White (4)	Other Race (5)	Female (6)	Male (7)	Low (8)	Normal or High (9)	Less Than High School (10)	High School or More (11)
NC Pre-K funding ($000s)	0.153**	0.184**	0.148**	0.122**	0.167**	0.154**	0.151**	0.167**	0.151**	0.155**	0.150**
	(0.004)	(0.011)	(0.015)	(0.009)	(0.006)	(0.004)	(0.004)	(0.005)	(0.004)	(0.010)	(0.005)
Observations	1,200,135	346,630	81,686	703,248	68,570	597,325	602,810	96,856	1,103,081	280,345	919,790
Enrollment mean	0.09	0.11	0.25	0.05	0.13	0.09	0.09	0.10	0.08	0.13	0.07
F statistic of instrument	1,462.28	264.64	93.19	183.08	708.17	1,237.15	1,473.82	977.27	1,403.33	226.55	763.84

Note. Standard errors are given in parentheses. Standard errors clustered at the county level. All models include Smart Start funding (in $000s), child and family covariates, county covariates, and county, program year, and quarter of birth fixed effects. See Table 3 for details.
NC Pre-K = North Carolina Pre-K.
$^{+}p < .10$.
$^{*}p < .05$.
$^{**}p < .01$.
$^{***}p < .001$.

specification with the key addition of NC Pre-K funding allocations, and the outcome is enrollment in NC Pre-K (1 = enrolled). Like our main results in Table 3, the first column includes the full sample of students in our study, and the remaining columns show results by subgroup. These coefficients represent percentage point changes in the probability of enrollment as a result of funding increases. The overall mean enrollment rates for each subgroup calculated across the entire study period are shown at the bottom of the table, as are the *F*-statistic tests of significance of the excluded instruments. The F-statistics are all well above 10, the required minimum threshold in the literature, and each of the NC Pre-K coefficients are significant, demonstrating that funding is meaningfully predictive of enrollment. Enrollment means also show differences across each subgroup in program participation.

We found that a $1,000 increase in county-level NC Pre-K funding increased the likelihood of enrollment in the entire population of children by 15.3 percentage points (pp; column 1). This effect was statistically significant, and fairly consistent, across each of our key subgroups as well. The results showing each covariate's prediction of NC Pre-K enrollment are also found in Table A3.

Local Average Treatment Effects for Children Enrolled in NC Pre-K (Second Stage)

Table 9 displays the local average treatment effects of NC Pre-K enrollment as a result of increases in county funding allocations (i.e., coefficients that are "local to" only those children who enrolled because of increased funding), on the 5th-grade academic composite. Across the full population of children in NC public schools, we found that enrollment in NC Pre-K improved a student's achievement by approximately 0.20 *SD*s at the end of 5th grade. This is larger than Deming's (2009) estimates of HS impacts on test scores at ages 7–10 (0.13 *SD*). The achievement effects appear to hold for both reading (coefficient = .22, $p < .01$) and math (coefficient = .16, $p < .05$) (Table A4).

Here, we found substantial heterogeneity in the relation between enrollment and achievement across population subgroups. Results in columns 3 and 10 show that the impacts of NC Pre-K attendance were most substantial for children who were Hispanic (0.30 *SD*s), and for children whose mothers had less than a high school degree (0.30 *SD*s). Effects were present but smaller in magnitude for children who were Black (0.18 *SD*s) or White (0.15 *SD*s; $p < .10$), and for children whose mothers had a high school degree or higher (0.14 *SD*s; $p < .10$). There was a nonsignificant local average treatment effect of NC Pre-K enrollment for children with low birthweight and children in the "other" race group (i.e., Native American, Asian, and mixed race students). As Table A4 reflects, achievement impacts tended to be larger for reading scores for most groups, though effects for Black and Hispanic children were quite similar in magnitude for both math and reading. Note that we were not able to test for significance across groups in the IV framework

TABLE 9

Instrumental Variables Results for the Effect of Enrollment on 5th Grade Academic Achievement

		Child's Race/Ethnicity				Gender		Birthweight		Mother Education	
	Full Sample (1)	Black (2)	Hispanic (3)	White (4)	Other Race (5)	Female (6)	Male (7)	Low (8)	Normal or High (9)	Less Than High School (10)	High School or More (11)
NC Pre-K enrollment	0.199**	0.177*	0.300**	0.146^{+}	0.071	0.205**	0.194*	0.149	0.204**	0.297**	0.142^{+}
	(0.073)	(0.086)	(0.096)	(0.075)	(0.116)	(0.075)	(0.076)	(0.104)	(0.071)	(0.079)	(0.076)
Observations	1,207,576	349,373	82,527	706,617	69,058	599,303	608,273	97,633	1,109,742	283,408	924,168

Note. Standard errors are given in parentheses. Standard errors clustered at the county level. All models include Smart Start funding (in $000s), child and family covariates, county covariates, and county, program year, and quarter of birth fixed effects. See Table 3 for details.
NC Pre-K = North Carolina Pre-K.
$^{+}p < .10$.
$^{*}p < .05$.
$^{**}p < .01$.
$^{***}p < .001$.

since each additional endogenous regressor requires an additional instrument, without which interactions of enrollment with each subgroup would not be identified. The standard errors of each subgroup estimate indicate overlap in the confidence intervals of the coefficients; thus, we cannot make strong statements about differential enrollment effects of NC Pre-K.

Sensitivity Checks

For these analyses, we incorporated student ED as a sensitivity exercise of the IV analyses. Here, we tested whether the population of children who were not economically disadvantaged, and thus were largely ineligible to enroll in NC Pre-K (with exceptions for children in military families and those with special needs), were affected by funding increases. If we found a "complier" effect for non-ED students in the IV analyses, this would work against the assumption that the instrument affected student outcomes only through enrollment. An important limitation of this sensitivity check is that our ED measure was derived from school records that indicated eligibility for free or reduced-price lunch (i.e., family income below 130% and 185% of the federal poverty level [FPL]). This differs from the NC Pre-K eligibility threshold (75% SMI). For a family of four in 2011, $41,348 was the income threshold for 185% FPL, but the SMI threshold was $50,975 (North Carolina Division of Child Development and Early Education, 2011). This nontrivial difference in family income eligibility means that the subgroup of students who met ED criteria was smaller than the actual population of students who were income-eligible for NC Pre-K, and some students in the non-ED subgroup were categorically eligible for NC Pre-K. Nonetheless, IV analyses of these subgroups can provide some additional validation of our approach.

We present the first- and second-stage IV results by ED status in Table A5. The first two columns present the first-stage results, which show that both ED and non-ED students enrolled in NC Pre-K. Non-ED students were eligible to enroll if they met a criterion of a disability or being a child of a military member. The effect of funding on enrollment was smaller for non-ED students than for the ED students. Importantly, the second stage results revealed no local average treatment effect of NC Pre-K enrollment on academic achievement for non-ED students (the estimated effect was nearly zero), and a significant 0.18 *SD* effect for ED students.

Summary of IV Analyses

IV analyses showed a strong effect of county-year funding on the likelihood that a child would receive a state-funded slot for pre-k, validating the premise of the primary approach to analysis. Further, we found that students who were more likely to enroll due to increases in state funding had higher levels of academic achievement in 5th grade by approximately 0.20 *SD*s, and these effects were consistently observed across demographic subgroups, as well as across math and reading scores. These findings support the hypotheses that state funding to a county in a year has impact on population-level outcomes by increasing the number of children who receive state-funded slots, and state-funded slots increase academic achievement among those receiving the slots.

VI. Examining Environmental Heterogeneity in the Long-Term Effects of NC Pre-K

In this chapter, we examine how the effects of NC Pre-K funding on the academic achievement composite vary depending on the other environments children encounter outside of pre-k. We replicated models on disaggregated measures of math and reading as reported in the Appendix. We focus on the achievement composite variable given the null main effect findings for special education and grade retention (with results for math and reading shown in the Appendix).

These models test our key hypotheses regarding complementarity, compensation, and additivity. For each environmental moderator, we examined whether the moderator produced a significant positive or negative main effect on children's academic achievement in 5th grade. We then examined whether the variable interacted with funding for pre-k using our two-way-fixed-effects model. As we detailed in Chapter III, if pre-k funding interacted positively with an environmental factor that also produced positive effects on children's learning, we would identify a complementary effect, whereby two positive environmental experiences interact to produce enhanced returns for children. Note that a negative interaction with a negative environmental experience would also be evidence for complementarity. Evidence for compensation would be found if we observed a positive interaction between a negative environmental exposure and pre-k funding, indicating that pre-k benefits children most in conditions where children encounter more environmental disadvantages. A compensatory effect would also be observed when an interaction between pre-k funding and a positive environmental exposure produces a negative coefficient (see Figures 1 and 2 in Chapter III). Additive effects would be observed if both environmental experiences (i.e., pre-k funding and the moderator measure) produce effects on children's achievement, but they do not significantly interact with one another.

Descriptive Results

We present descriptive statistics for each of our moderators in Table 10 corresponding to 2003, the first year of full NC Pre-K implementation, for all children in our analytic sample. The values in the first column are for all students in 5th grade in 2003. The remaining three columns present descriptive values by the NC Pre-K funding tercile of the students' county, whereby tercile 1 captures the counties with the lowest levels of NC Pre-K

TABLE 10

Descriptive Statistics of Moderator Variables by Terciles of County NC Pre-K Funding (2003 Program Year)

	All Counties *M* (*SD*) (1)	Tercile 1 of NC Pre-K Funding *M* (*SD*) (2)	Tercile 2 of NC Pre-K Funding *M* (*SD*) (3)	Tercile 3 of NC Pre-K Funding *M* (*SD*) (4)
NC Pre-K funding ($000s; no zero values)	0.32	0.15	0.27	0.53
	(0.21)	(0.05)	(0.03)	(0.21)
Alternative ECE services				
Head Start 4-year-old saturation (%)	11.68	9.52	11.06	14.76
	(6.42)	(6.08)	(3.88)	(7.27)
Early Head Start 0–3-year-old saturation (%)	0.35	0.27	0.26	0.52
	(1.39)	(0.50)	(0.67)	(2.27)
Smart Start funding ($000s)	2.38	2.25	2.28	2.60
	(0.78)	(0.66)	(0.86)	(0.80)
HS grantee in county	0.74	0.84	0.80	0.58
	(0.44)	(0.37)	(0.40)	(0.49)
2003 HS 4-year-old enrollment (%)	10.33	8.66	10.74	11.97
	(4.98)	(4.08)	(3.46)	(6.26)
School characteristics				
School-average achievement composite (lagged;std.)	0.03	0.18	0.03	−0.15
	(0.99)	(1.01)	(0.95)	(0.97)
Local PPE (in 2019 $000s)	2.33	2.46	2.45	2.09
	(0.75)	(0.63)	(0.76)	(0.80)
State PPE (in 2019 $000s)	6.49	6.33	6.41	6.76
	(0.69)	(0.52)	(0.48)	(0.91)
Federal PPE (in 2019 $000s)	1.26	1.08	1.19	1.53
	(0.46)	(0.38)	(0.35)	(0.50)

(*Continued*)

TABLE 10. (*Continued*)

	All Counties *M* (*SD*) (1)	Tercile 1 of NC Pre-K Funding *M* (*SD*) (2)	Tercile 2 of NC Pre-K Funding *M* (*SD*) (3)	Tercile 3 of NC Pre-K Funding *M* (*SD*) (4)
Teacher characteristics				
Student–teacher ratio in the school	15.01	14.99	14.91	15.14
	(2.05)	(2.02)	(2.06)	(2.07)
National Board Certified teachers (%)	13.28	13.32	13.39	13.15
	(8.78)	(8.41)	(8.97)	(9.04)
Teachers with <3 years experience (%)	20.94	21.27	21.42	20.14
	(10.59)	(10.24)	(10.87)	(10.72)
Annual teacher turnover (%)	11.17	10.97	11.21	11.39
	(6.40)	(6.08)	(6.22)	(6.90)
County economic factors (in the year of the test)				
Job loss (% affected)	1.09	1.13	1.34	0.83
	(0.80)	(0.80)	(0.85)	(0.68)
Median family income ($00s)	679.13	743.76	673.50	607.12
	(117.23)	(113.29)	(85.27)	(100.29)
SNAP recipients (%)	13.04	10.88	12.84	15.77
	(4.64)	(4.12)	(2.90)	(5.00)
Medicaid recipients (%)	17.97	15.18	17.31	21.82
	(5.86)	(5.07)	(3.86)	(6.02)
Observations	71,017	27,892	19,629	23,496

Note. ECE = early childhood education; HS = Head Start; NC Pre-K = North Carolina Pre-K; PPE = per-pupil expenditure; *SD* = standard deviation; SNAP = Supplemental Nutrition Assistance Program.

funding and tercile 3 contains the counties with the highest levels of funding. These latter columns allow one to see relations between our key treatment variable, NC Pre-K funding, and each of our moderators.

As a means-tested program, greater NC Pre-K funding should be allocated to communities with higher levels of poverty and economic need. This is borne out in the data, whereby each moderator that corresponds with ED increases as NC Pre-K funding increases (e.g., HS saturation, state and federal PPE, % SNAP recipients, % Medicaid recipients). Conversely, moderators that indicate economic advantage decrease as NC Pre-K funding increases (e.g., school academic achievement, local PPE, median family income).

Also important for our methodological approach is that there exists variation in both the treatment and the moderator variables to identify relations. We selected three moderators from each of the three environmental contexts we examine (early childhood, school, county economics) to illustrate the variation in both NC Pre-K funding and moderators. Figure 4 displays scatterplots of county HS enrollment, school average achievement, and county job loss with NC Pre-K funding. The figures clearly illustrate that our moderators vary along the continuum of county-level NC Pre-K funding allocations.

We also examined the relations between pre-k funding exposure and our moderators with regression analyses. Table 11 summarizes the results from

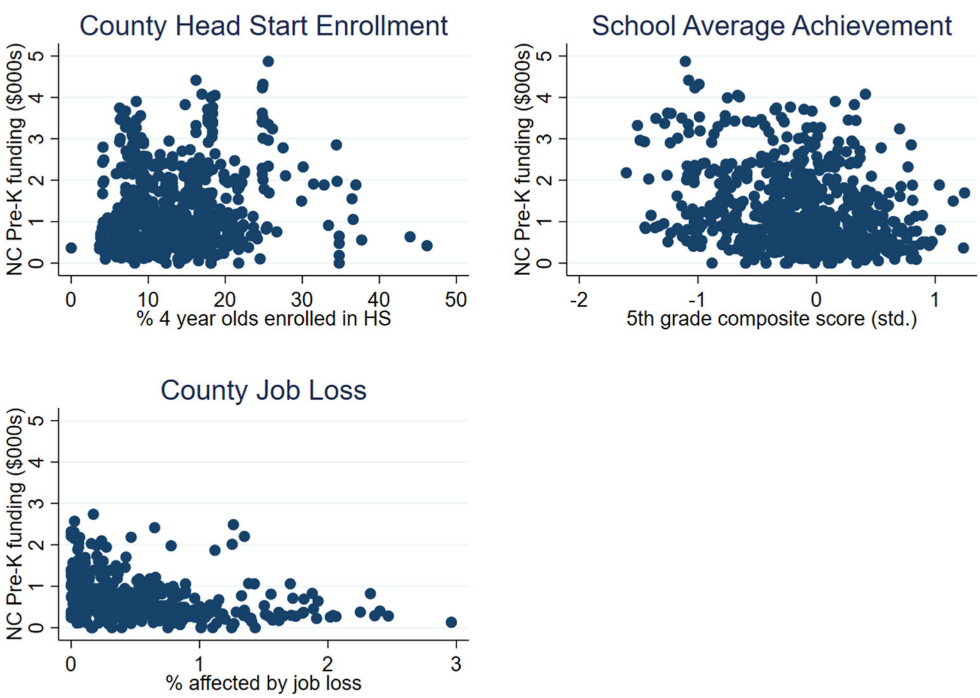


FIGURE 4.—Scatterplot of selected moderator variables and North Carolina Pre-K (NC Pre-K) Funding. HS = Head Start.

TABLE 11
Endogeneity Checks for Key Moderator Variables

	NC Pre-K Funding
Alternative ECE services	
HS Saturation (in 10%)	0.134*
	(0.062)
Max EHS Enrollment (in 10%)	0.011
	(0.010)
Smart Start funding ($000s)	−0.038
	(0.065)
School characteristics	
School-average achievement composite (lagged;std.)	0.006
	(0.026)
Local PPE (in 2019 $000s)	0.043+
	(0.024)
State PPE (in 2019 $000s)	0.056*
	(0.027)
Federal PPE (in 2019 $000s)	0.004
	(0.015)
Teacher characteristics	
Student–teacher ratio in the school	0.234*
	(0.091)
National Board Certified teachers (in 10%)	−0.064+
	(0.038)
Teachers with <3 years experience (in 10%)	0.076
	(0.054)
Annual teacher turnover (in 10%)	0.046
	(0.032)
County economic factors	
Job loss (% affected)	0.035
	(0.076)
Median family income	10.918+
	(5.610)
SNAP recipients (%)	0.635**
	(0.233)
Medicaid recipients (%)	−0.397*
	(0.188)

Note. Each coefficient and standard error was generated from a county-by-cohort level regression model in which a given moderator variable was regressed on NC Pre-K funding with county and cohort fixed effect included. All but one regression had $n = 1{,}800$ observations. The Job Loss regression model had 1,468 observations.
ECE = early childhood education; EHS = Early Head Start; NC Pre-K = North Carolina Pre-K; PPE, per-pupil expenditure; SNAP = Supplemental Nutrition Assistance Program.
$^{+}p < .10$.
$^{*}p < .05$.
$^{**}p < .01$.
$^{***}p < .001$.

regressing each of the moderator variables, aggregated to the county-year level, on NC Pre-K funding at the county-year level, including county and year fixed effects with standard errors clustered by county. Across the 15 moderators, eight of the measures are at least marginally significantly correlated with NC Pre-K funding ($p < .10$). We consider our moderators to be endogenous and do not endorse causal interpretations of the relations we report here.

Next, we present the moderation results for each of the three conceptual moderator groups (ECE, school environment [classroom and teacher factors], and economic environment) in relation to the 5th-grade achievement composite measure. Our tables display the coefficients for the key ingredients necessary for assessing the theories of complementarity, compensatory effects, and additive effects: (1) main effect of NC Pre-K funding, (2) the main effect of moderator, and (3) the interaction between NC Pre-K funding and each moderator. To provide evidence for either complementarity or compensation, we argue that a given moderator must produce a statistically significant main effect on achievement and a statistically significant interaction term with NC Pre-K exposure, as both terms are essential for understanding the theoretical implications of the interaction. For additive effects, we must observe a statistically significant effect of the hypothesized moderator, with a null interaction term between the moderator and NC Pre-K exposure.

Other covariates and fixed effects included in each model are omitted from the display. In each table, the first model specification omits the moderators and interactions, with only the NC Pre-K funding main effect coefficient displayed, to examine any differences between our results in Chapter V and the moderator analysis subsample (i.e., observations with nonmissing moderator measures). For each moderator, we present first a specification including only the moderator's main effect, and then a second specification that includes the interaction of that moderator with NC Pre-K funding. The final omnibus specification includes all moderators and interaction terms together and tests their joint significance using a post hoc F-test. Moderation analyses of reading and math achievement outcomes measured separately are available for each moderator group in Tables A6–A13. Each moderator that is a continuous variable was centered at the variable's mean to ease interpretation, such that the main effect of NC pre-K funding in each interaction model represents the change in the outcome at the moderator's mean value.

Access to Other Early Childhood Programs

Moderation results for the early childhood educational environment are shown in Table 12. The coefficient in column 1 demonstrates that the main effect of NC Pre-K funding for the ECE moderator analytic subsample is consistent with our Chapter V results. Columns 2 and 3 test for moderation by HS saturation in a county during children's age-4 year. These results show that the association between greater HS saturation and increases in student achievement was not statistically significant. Further, the interaction between

TABLE 12

Effects of NC Pre-K Funding on Academic Achievement: Moderation by the Early Childhood Environment

		Moderator: 4-Year-Old HS Enrollment		Moderator: EHS Enrollment (Ages 0–3)		Moderator: Smart Start			HS as a "springboard" for NC Pre-K	
									Mod: HS Indicator	Mod: 2003 HS Enrollment
	Baseline (1)	Main Effect (2)	Int. (3)	Main Effect (4)	Int. (5)	Main Effect (6)	Int. (7)	All Mods. (8)	Int. Only (9)	Int. Only (10)
NC Pre-K funding ($000s)	0.030**	0.030**	0.038**	0.031**	0.033**	0.030**	0.021^{+}	0.029*	0.032***	0.031**
	(0.011)	(0.011)	(0.012)	(0.011)	(0.011)	(0.011)	(0.013)	(0.014)	(0.009)	(0.012)
HS Saturation (in 10pp)		−0.007	−0.004					−0.005		
		(0.007)	(0.007)					(0.007)		
HS Saturation (in 10pp) × NC Pre-K funding			−0.013					−0.008		
			(0.008)					(0.009)		
Max EHS Enrollment (in 10pp)				−0.080*	−0.061^{+}			−0.061^{+}		
				(0.031)	(0.032)			(0.032)		
Max EHS Enrollment (in 10pp) × NC Pre-K funding					−0.028			−0.020		
					(0.023)			(0.025)		
Smart Start funding ($000s)						0.002	0.002	0.001		
						(0.006)	(0.006)	(0.006)		
Smart Start funding ($000s) × NC Pre-K funding							0.012	0.010		
							(0.010)	(0.010)		
HS grantee in county × NC Pre-K funding									−0.003	
									(0.013)	

(Continued)

TABLE 12. (*Continued*)

		Moderator: 4-Year-Old HS Enrollment		Moderator: EHS Enrollment (Ages 0–3)		Moderator: Smart Start			HS as a "springboard" for NC Pre-K	
									Mod: HS Indicator	Mod: 2003 HS Enrollment
	Baseline (1)	Main Effect (2)	Int. (3)	Main Effect (4)	Int. (5)	Main Effect (6)	Int. (7)	All Mods. (8)	Int. Only (9)	Int. Only (10)
2003 HS 4-year-old enrollment × NC Pre-K funding										−0.000 (0.001)
Observations	1,207,576	1,207,576	1,207,576	1,207,576	1,207,576	1,207,576	1,207,576	1,207,576	1,207,576	1,207,576
p Value (interactions = 0)								.380		

Note. Standard errors clustered at the county level. All models include Smart Start funding (in $000s), child and family covariates, county covariates, and county, program year, and quarter of birth fixed effects. See Table 3 for details. No main effect models or main effects are shown for time-invariant measures of 2003 HS enrollment and the HS grantee in county indicator; these variables drop out in county fixed effects regression and can only be recovered through their interaction with a time-varying county feature (i.e., NC Pre-K funding).
EHS = Early Head Start; HS = Head Start; NC Pre-K = North Carolina Pre-K.
$^{+}p < .10$.
$^{*}p < .05$.
$^{**}p < .01$.
$^{***}p < .001$.

HS and NC Pre-K was also not significant. The NC Pre-K effect remained significant, indicating that HS and NC Pre-K funding effects were independent.

Our next moderator measures capture access to early childhood programs that largely target services to children before they become age eligible for NC Pre-K: EHS and Smart Start. The main effect of EHS saturation (the maximum value of county enrollment across all EHS age groups) was negative and significant (−0.08 *SD*; column 4) likely reflecting a community's ED. The interaction between EHS and NC Pre-K was not significant, indicating support for additive effects of these two variables (column 5). The Smart Start coefficient was 0.002 *SD* but was not significant (column 6), and its interaction with NC Pre-K was positive but similarly not significant.

On the right-hand side of the table are our results testing the possibility that HS counties, defined at the start of statewide NC Pre-K implementation in 2003, were more “shovel-ready” to implement NC Pre-K, providing a “springboard” for successful program scale-up. Our two measures, an indicator for HS presence in a county and the HS saturation level in 2003, were included here as moderators. Note that we cannot show main effect models or main effect coefficients for these moderators because they are time-invariant, and therefore drop out with county fixed effects included in the model. Neither of these measures' interactions with NC Pre-K were significant, suggesting that we found no support for the “shovel ready” hypothesis.

School and Teacher Environments

School Quality

Moderation results for school measures are shown in Table 13. The coefficient in Column 1 indicates the main effect of NC Pre-K funding for the school moderator analytic subsample is significant and consistent with our Chapter V results. The first moderator analysis, beginning in column 2, is the lagged measure of the school-level academic composite of test scores (standardized reading and math averaged, then restandardized). Here we observed a positive, significant, and fairly large main effect of this measure (0.18 *SD*). The interaction of school achievement and NC Pre-K funding was also significant but negative, and approximately one half the size of the NC Pre-K coefficient in this model (−0.01 *SD*). This pattern suggests that the positive effect of pre-k exposure was attenuated in higher performing schools. This also suggests that students who attended lower-achieving elementary schools benefitted more from exposure to funding for pre-k, providing evidence for a compensatory effect.

Federal PPEs had a significant negative main effect on students' academic composite scores (−0.11 *SD*). Federal funding for schools is allocated based

TABLE 13

Effects of NC Pre-K Funding on Academic Achievement: Moderation by Elementary School Characteristics

	Baseline With Chapter Subsample	Moderator: (Lagged) Academic Composite		Moderator: Local PPE		Moderator: State PPE		Moderator: Federal PPE		
		Main Effect	Interaction	Main Effect	Interaction	Main Effect	Interaction	Main Effect	Interaction	All Mods.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
NC Pre-K funding ($000s)	0.029**	0.023**	0.019*	0.029**	0.028*	0.030**	0.030*	0.028**	0.013	0.004
	(0.011)	(0.008)	(0.008)	(0.011)	(0.011)	(0.011)	(0.012)	(0.010)	(0.011)	(0.009)
School-average achievement composite (lagged;std.)		0.183***	0.187***							0.184***
		(0.003)	(0.003)							(0.003)
School-average achievement composite × NC Pre-K funding			−0.008*							−0.003
			(0.003)							(0.003)
Local PPE (in 2019 $000s)				0.058***	0.059***					0.023***
				(0.007)	(0.008)					(0.007)
Local PPE × NC Pre-K funding					−0.003					0.000
					(0.009)					(0.008)
State PPE (in 2019 $000s)						−0.029***	−0.028**			0.034***
						(0.008)	(0.008)			(0.008)

(Continued)

TABLE 13. (*Continued*)

	Baseline With Chapter Subsample	Moderator: (Lagged) Academic Composite		Moderator: Local PPE		Moderator: State PPE		Moderator: Federal PPE		
		Main Effect	Interaction	Main Effect	Interaction	Main Effect	Interaction	Main Effect	Interaction	All Mods.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
State PPE × NC Pre-K funding							−0.001			−0.017*
							(0.007)			(0.007)
Federal PPE (in 2019 $000s)								−0.110***	−0.128***	−0.030
								(0.020)	(0.022)	(0.021)
Federal PPE × NC Pre-K funding									0.032*	0.053**
									(0.014)	(0.016)
Observations	1,149,113	1,149,113	1,149,113	1,149,113	1,149,113	1,149,113	1,149,113	1,149,113	1,149,113	1,149,113
p Value (interactions = 0)										.031

Note. Standard errors are given in parentheses. Standard errors clustered at the county level. All models include Smart Start funding (in $000s), child and family covariates, county covariates, and county, program year, and quarter fixed effects. See Table 3 for details.
NC Pre-K = North Carolina Pre-K; PPE = per-pupil expenditure.
$^{+}p < .10$.
$^{*}p < .05$.
$^{**}p < .01$.
$^{***}p < .001$.

on community ED (federal PPE is correlated with county SNAP and Medicaid recipients at .74; Table A2), whereby more resources are targeted at higher poverty schools. We found a positive and significant interaction between federal PPE and NC Pre-K funding (0.03 *SD*). Because federal PPE likely represents a measure of disadvantage in our models, this interaction likely represents a compensatory effect, as NC Pre-K exposure protected students from future difficult educational conditions. This pattern remains stable in the omnibus model (column 10).

State PPEs had a significant negative main effect on students' academic composite scores (−0.03 *SD*). State funding policy provided an equal amount of funding per enrolled child, plus extra funding for children in need. This variable did not interact significantly with pre-k funding, indicating additive effects.

Local PPE was significantly and positively associated with achievement (0.06 *SD*), perhaps because local PPE reflects the wealth of the community as assessed through property taxes. This variable did not interact significantly with pre-k funding, indicating additive effects.

Teacher Characteristics

Results for moderation by teacher characteristics (measured at the school-level) are shown in Table 14. Results in Column 1 indicate the main effect of NC Pre-K funding for the teacher moderator analytic subsample was significant and consistent with the Chapter V results. We began with student–teacher ratio, where we observed a significant positive main effect (0.01 *SD*, Column 2), but no significant interaction with NC Pre-K funding. We expected a *negative* coefficient on student–teacher ratio, and probed this result with the inclusion of additional school covariates in the robustness section, where we found the expected negative relation between student–teacher ratio and achievement (−0.003 *SD*, $p < .05$; Table A17). Thus, this finding likely represents unobserved school-level confounding. In our main analyses, we did not observe a significant interaction effect, indicating that pre-k funding and student–teacher ratio operated as additive effects.

Results for the proportion of teachers who are National Board Certified indicated a significant positive main effect on student achievement (0.05 *SD*, column 4). A nonsignificant interaction effect indicated that pre-k funding and this variable operated as additive effects.

The proportion of teachers who are inexperienced (<3 years' experience) was significantly and negatively associated with student achievement (−0.05 *SD* [column 6]). We found a significant positive interaction between this variable and NC Pre-K funding (0.01 *SD*), such that the NC Pre-K funding effect was stronger if children later attended elementary schools with less experienced teachers. This pattern again suggested a compensatory effect,

TABLE 14

Effects of NC Pre-K Funding on Academic Achievement: Moderation by Teacher Characteristics

	Baseline With Chapter Subsample	Moderator: Student–Teacher Ratio		Moderator: National Board Certification		Moderator: Teacher Experience		Moderator: Teacher Turnover		
		Main Effect	Interaction	Main Effect	Interaction	Main Effect	Interaction	Main Effect	Interaction	All Mods.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
NC Pre-K funding ($000s)	0.029**	0.029**	0.029*	0.031**	0.031**	0.027**	0.026*	0.028*	0.028*	0.027*
	(0.010)	(0.011)	(0.011)	(0.010)	(0.010)	(0.010)	(0.010)	(0.011)	(0.011)	(0.011)
Student–teacher ratio in the school		0.008***	0.008***							0.006***
		(0.001)	(0.001)							(0.001)
Student–teacher ratio in the school × NC Pre-K funding			0.003							0.001
			(0.003)							(0.002)
National Board Certified teachers (in 10pp units)				0.045***	0.045***					0.024***
				(0.006)	(0.008)					(0.006)
National Board Certified teachers × NC Pre-K funding					0.001					0.006
					(0.006)					(0.006)
Teachers with <3 years experience (in 10pp units)						−0.047***	−0.051***			−0.042***
						(0.003)	(0.004)			(0.004)
Teachers with <3 years							0.008*			0.011**
							(0.004)			(0.003)

(Continued)

TABLE 14. (*Continued*)

	Baseline With Chapter Subsample (1)	Moderator: Student–Teacher Ratio Main Effect (2)	Moderator: Student–Teacher Ratio Interaction (3)	Moderator: National Board Certification Main Effect (4)	Moderator: National Board Certification Interaction (5)	Moderator: Teacher Experience Main Effect (6)	Moderator: Teacher Experience Interaction (7)	Moderator: Teacher Turnover Main Effect (8)	Moderator: Teacher Turnover Interaction (9)	All Mods. (10)
experience × NC Pre-K funding										
Annual teacher turnover (in 10pp units)								−0.034*** (0.003)	−0.035*** (0.003)	−0.014*** (0.002)
Annual teacher turnover × NC Pre-K funding									0.001 (0.003)	0.002 (0.003)
Observations	1,187,543	1,187,543	1,187,543	1,187,543	1,187,543	1,187,543	1,187,543	1,187,543	1,187,543	1,187,543
p Value (interactions = 0)										.013

Note. Standard errors are given in parentheses. Standard errors clustered at the county level. All models include Smart Start funding (in $000s), child and family covariates, county covariates, school-district covariates, and county, program year, and quarter fixed effects. See Table 3 for details. School district covariates: federal, state, and local per pupil expenditures in the school district.
NC Pre-K = North Carolina Pre-K.
$^{+}p < .10$.
$^{*}p < .05$.
$^{**}p < .01$.
$^{***}p < .001$.

indicating that NC Pre-K funding protected children from the adverse effect of attending a school with a high proportion of inexperienced teachers.

Teacher turnover (proportion of teachers who leave the school annually) was significantly and negatively associated with academic achievement (−0.03 *SD* [column 8]). We did not find a significant interaction between NC Pre-K funding and teacher turnover, indicating these variables operated as additive effects.

County Economic Conditions

Our final set of moderators captured the economic conditions of the county during the child's 5th grade year. Because these measures were primarily negatively valanced (with the exception of median family income), these models tested whether exposure to NC Pre-K was compensatory by protecting children against future adverse environmental economic conditions.

Job Loss

This set of analyses builds on work by Ananat and colleagues (2011), which examined the effect of county-level job losses in North Carolina on student test scores. As a consequence of widespread plant closings in the state (as the push for globalization drove manufacturing jobs overseas) the authors found that job loss shocks occurring immediately before students took standardized tests lowered student reading test score performance in 8th grade. The job loss effects were most strongly felt by students who came from non-college-educated parents. They also reported negative effects of job losses on 4th grade reading scores that were not as robustly significant. Table A14 reports year-by-year descriptive statistics of job losses, and Appendix Figure A1 illustrates the variation in job losses across all 100 counties in several representative years. These numbers indicated that job losses fluctuated from year to year, with average job loss across counties in a given year ranging from approximately 0.58–1.43 percentage points. We did not observe statistically significant correlations between the county-level job losses and NC Pre-K funding (Table 11)[7] suggesting that negative shocks to county employment conditions are exogenous to NC Pre-K funding.

Results for job loss are presented in Table 15. Because our data on county-level employment was only available through 2011, these models include a smaller sample of 824,416 students who reached Grade 5 by the 2010-2011 academic year (i.e., only earlier cohorts of children exposed to NC Pre-K were available for this analysis). In Column 1, we present the main effect of NC Pre-K for the reduced sample of children with nonmissing job loss data and find a much larger main effect of NC Pre-K funding than the main result reported in Chapter V (0.06 *SD*). This larger main effect most

TABLE 15

Effects of NC Pre-K Funding on Academic Achievement: Moderation by County-Level Job Losses

	Baseline With Chapter Subsample	Moderator: County-Level Job Loss					
		Baseline	Main Effect	Interaction	Baseline	Main Effect	Interaction
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
NC Pre-K funding ($000s)	0.059*	0.056*	0.056*	0.061*	0.008	0.008	0.012
	(0.027)	(0.027)	(0.027)	(0.025)	(0.019)	(0.019)	(0.019)
Percent affected by job loss per full academic year			0.003	0.002		0.004	0.003
			(0.003)	(0.003)		(0.004)	(0.003)
Percent affected by job loss per full academic year × NC Pre-K funding				0.017			0.029*
				(0.019)			(0.013)
Analysis specific controls							
Year of test FE		X	X	X	X	X	X
Time trends and quadratic					X	X	X
Observations	824,416	824,416	824,416	824,416	824,416	824,416	824,416

Note. Standard errors are given in parentheses. Standard errors clustered at the county level. All models include Smart Start funding (in $000s), child and family covariates, county covariates, and county, program year, and quarter fixed effects. See Table 3 for details.

NC Pre-K = North Carolina Pre-K.

$^{+}p < .10$.

$^{*}p < .05$.

$^{**}p < .01$.

$^{***}p < .001$.

likely reflects the aforementioned issue that variation in state funding for later cohorts of students was smaller; eliminating these later cohorts increased the effect size for NC Pre-K funding. In Column 2, we estimate the NC Pre-K main effect in our analysis subsample but include here a fixed effect for year of test (not included in our other moderation analyses). Because job losses were reasonably exogenous to other factors that might affect student test scores and NC Pre-K funding, we control for any year-to-year differences that might have occurred alongside negative labor market shocks during the year of the test (e.g., the Great Recession). Here, the main effect of pre-k was significant and positive (0.06 *SD*; $p < .05$).

We then add the job loss measure to the model in Column 3, which did not significantly predict student test scores. Column 4 includes the interaction between the job loss and NC Pre-K funding, which was not significant. In Columns 5 through 7, we test an additional specification that adds county linear and quadratic time trends to the model, which more closely match the specification used by Ananat and colleagues (2011). Here, we did not observe significant main effects of NC Pre-K or the job loss measure, but did find a positive and significant interaction between job loss shocks and funding for NC Pre-K (0.03 *SD*). If we assume, based on the work of Ananat and colleagues, that job losses are not positive economic shocks, then these findings may suggest another compensatory effect, whereby NC Pre-K effects on achievement were enhanced in counties that experienced worse job losses. However, the lack of significant main effect for job losses in our model seriously tempers any clear theoretical interpretation of these findings.

Economic Indicators

Our next set of moderators more broadly measured the overall economic status of a county: median family income, the proportion of residents receiving SNAP benefits, and the proportion of residents enrolled in Medicaid. Unlike our job loss measure, these other economic indicators do not represent exogenous shocks to a child's environment. Indeed, as we show in Table 11, all three of these indicators were significantly correlated with NC Pre-K funding and should be interpreted cautiously.

The economic indicator data are available for our entire panel, and therefore we include nearly the full sample of NC children considered in our main effect analyses presented in Chapter V (i.e., $n = 1{,}207{,}336$). In Table 16, column 1 indicates the main effect of NC Pre-K funding on student achievement for this analytic sample was significant and consistent in magnitude with the findings reported in Chapter V (0.03 *SD*). We then added to the model county-level median family income, which significantly predicted higher child test scores (0.001 *SD*). We did not find a significant interaction with NC Pre-K funding (column 3), indicating that pre-k funding and median family income operate as additive effects.

TABLE 16

EFFECTS OF NC PRE-K FUNDING ON ACADEMIC ACHIEVEMENT: MODERATION BY COUNTY-LEVEL ECONOMIC FACTORS (IN THE YEAR OF THE TEST)

	Baseline With Chapter Subsample	Moderator: Median Family Income		Moderator: % SNAP Recipients		Moderator: % Medicaid Recipients		
		Main Effect	Interaction	Main Effect	Interaction	Main Effect	Interaction	All Mods.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
NC Pre-K funding ($000s)	0.030**	0.028*	0.008	0.045***	−0.002	0.031*	0.026	−0.008
	(0.011)	(0.011)	(0.017)	(0.013)	(0.021)	(0.013)	(0.018)	(0.025)
Estimated median family income ($00s)		0.001***	0.001***					0.000^{+}
		(0.000)	(0.000)					(0.000)
Estimated median family income × NC Pre-K funding			−0.000					−0.000
			(0.000)					(0.000)
SNAP recipients (%)				−0.018***	−0.022***			−0.020***
				(0.004)	(0.005)			(0.006)
SNAP recipients × NC Pre-K funding					0.004*			0.009**
					(0.002)			(0.003)
Medicaid recipients (%)						−0.025***	−0.025***	−0.008*
						(0.003)	(0.003)	(0.003)
Medicaid recipients × NC Pre-K funding							0.000	-0.005^{+}
							(0.001)	(0.003)
Observations	1,207,336	1,207,336	1,207,336	1,207,336	1,207,336	1,207,336	1,207,336	1,207,336
p Value (interactions = 0)								.022

Note. Standard errors are given in parentheses. Standard errors clustered at the county level. All models include Smart Start funding (in $000s), child and family covariates, county covariates, and county, program year, and quarter fixed effects. See Table 3 for details.
NC Pre-K = North Carolina Pre-K; SNAP = Supplemental Nutrition Assistance Program.
$^{+}p < .10$.
$^{*}p < .05$.
$^{**}p < .01$.
$^{***}p < .001$.

Next, we found that the county proportion of SNAP recipients was significantly and negatively associated with student achievement (−0.02 *SD*). In Column 5, we added the interaction with NC Pre-K funding and found a significant positive interaction, indicating that NC Pre-K effects were enhanced in counties that had more residents qualify for SNAP benefits (0.004 *SD*). This pattern provides support for the compensatory hypothesis that exposure to pre-k protects children from the later adverse effect of attending school in a community with a high rate of SNAP receipt.

We found a significant negative main effect of the county proportion of Medicaid enrollees on student achievement, indicating lower scores in counties with higher rates of Medicaid enrollment. We did not find a significant interaction between funding for pre-k and Medicaid, indicating that pre-k and Medicaid rates operated as additive effects.

Results for Reading and Math

We also estimated each of our moderation analyses separately by reading and math test outcomes to assess whether our findings differed across outcome domains, shown in Tables A6–A13. Here, we report instances where a significant moderation finding on the academic composite outcome measure appeared to have a stronger impact on reading or math achievement. For school achievement, federal PPE, and percentage in the county on SNAP, moderation results were statistically significant and similar in magnitude for both math and reading. For the interaction that we observed on the composite score for teacher experience, we found that this interaction was driven by effects on math achievement, as the interaction for reading achievement was null and close to zero in magnitude. The job loss interaction observed for the model that included quadratic time trends was also only observed for math, and not for reading.

Sensitivity Tests

In each of our moderator groups, we estimated different specifications of the moderator analyses and variations of the construct measures to test the sensitivity of the findings presented here. For ECE moderation we used different measures of HS (derived from PIR data only; averaged values of PIR-only and HS Cluster calculations of enrollment for 4-year-olds, 3-year-old enrollment instead of 4-year-old enrollment; Table A15) and EHS (mean county EHS enrollment of 0–3-year-olds instead of max county EHS enrollment). When HS enrollment was derived strictly from the PIR data (i.e., assuming the county in which the HS grantee is located is the county where children are served; no cross-county service delivery), our county-level saturation values were consistently lower than the values derived using the cluster methodology (Appendix B). Both measures involve assumptions about service delivery within and across counties over time, and we take the view that the "true" value of county HS saturation is likely bounded by these two calculations. In turn, we estimated our moderation analyses with both the PIR-only value, and an average of the two values. For the PIR-only measure, we found a small, negative, but nonsignificant main effect of HS, and a significant and positive interaction with NC Pre-K funding. However, our findings followed a similar pattern as that shown in Table 12 when we used the averaged value of 4-year-old enrollment from the two methodologies, finding a nonsignificant but negative main effect of HS enrollment, and a negative interaction with NC Pre-K that was not significant. We do not find any main effect or moderation evidence from the age-3 HS enrollment moderator measure (i.e., measured as the age-3 HS saturation in the year prior to the child's age-eligible pre-k year). For EHS, using the mean county enrollment of 0–3-year-olds rather than the maximum produced the same pattern of results, with a negative and significant main effect, but no interaction with NC Pre-K funding.

For our school and teacher moderation analyses, we tested specifications that added school-level demographic characteristics of the students (% Black, % Hispanic, % EDS, average daily membership). For most specifications these controls did not meaningfully change the results (Tables A16 and A17). However, the main effect of student–teacher ratio became negative, and produced a marginally significant and negative interaction with NC Pre-K funding, which aligns with our hypotheses about this measure's relation with achievement. The addition of school demographic controls also changed the coefficient on federal PPE to nonsignificance, probably due to its correlation with the proportion of students who were economically disadvantaged. We created an alternative school achievement measure that averaged a school's achievement composite score across Grades 3 through 5, rather than Grade 5 only, which resulted in a slightly larger interaction coefficient between school achievement and NC Pre-K funding (Table A18).

In the county economic conditions analyses, we tested for interactions between our three economic indicators measured at two different time points: the child's birth year (Table A19), and at age 4 (year of expected pre-k enrollment based on birthdate; Table A20). We did not observe any statistically significant interactions.

Intersectional Moderation: Crossing Environmental Moderators With Child and Family Characteristics

Following the intersectionality analyses presented in Chapter V, we further probed whether the four observed significant interaction effects (for school-average achievement, federal PPE, percent of teachers with below 3 years of experience, and the percent of a county's population receiving

SNAP) were consistent across key subgroups based on child and family characteristics. Each moderator is presented as a separate panel on Table 17, with results for each subgroup shown in columns 2–9, and full sample results (i.e., coefficients reported in sections above) shown in the first column for comparison. Key here are any subgroup interaction coefficients that differ in magnitude or direction from the full sample results.

For the three environmental measures indicating adverse environments (federal PPE, teachers with <3 years of experience, and percent of SNAP recipients in a county), we found the main effect of these variables and their interaction effects did not substantially differ across key population subgroups. Interaction effects were not significant for every group, but interaction coefficients were typically in the same direction and similar in magnitude (though these results were far noisier than those estimated for the full sample).

We found a different pattern for average school achievement. For Black students, attending a school with higher average achievement appeared to provide a complementary relation with NC Pre-K funding. In stark contrast to the interaction coefficient for the full sample (−0.01), the interaction term for Black students was significant in the *positive* direction (0.02, $p < .001$). Given the apparent difference in the school achievement interaction for Black students, we pursued several post-hoc analyses to better understand this effect.

Post-hoc descriptive analyses shown in Figure 5 display the distributions of the school average achievement composite for the full sample of children and for the sample of children who were Black, both on the same scale. The leftward skew of the average school achievement for the schools in which Black children attend demonstrates the legacy of systemic racism and the lack of investment in Black communities. It is clear that Black children in NC were more likely to attend a school with low average achievement.

Distributional differences by subgroups indeed undergird our apparently discrepant regression results. Based on this descriptive finding, we then estimated the school achievement moderation analyses for Black students for each of three different segments of the distribution: >1 *SD* below the statewide mean of school average achievement, between −1 and 1 *SD*s around the statewide mean, and >1 *SD* above the statewide mean. In other words, we examined subgroup models similar to those shown in Chapter V, but focused the sample on only Black students, and further split the sample into subgroups based on school-level performance. Observing the magnitude and direction of the coefficient for NC Pre-K funding across each of these subgroup models can help identify how the positive interaction between school achievement and pre-k funding for Black students operates across the distribution of elementary school performance.

Shown in Table A21, these regression results reveal the complementary relation between school average achievement and NC Pre-K funding for Black students did not hold for students attending schools with higher

TABLE 17

EFFECTS OF NC PRE-K FUNDING ON ACADEMIC ACHIEVEMENT: INTERSECTIONAL MODERATION BY CHILD, FAMILY AND ENVIRONMENTAL FACTORS

		Race/Ethnicity				Mother Education		Birthweight	
	Full Sample (1)	Black (2)	Hispanic (3)	White (4)	Other Races (5)	Less Than High School (6)	High School or More (7)	Low (8)	Normal or High (9)
School-average achievement composite (lagged;std.)									
NC Pre-K funding ($000s)	0.019*	0.039**	0.044***	0.012^{+}	0.018	0.040***	0.012	0.021	0.018*
	(0.008)	(0.013)	(0.012)	(0.007)	(0.013)	(0.011)	(0.008)	(0.015)	(0.008)
School-average achievement composite (lagged;std.)	0.187***	0.138***	0.144***	0.205***	0.208***	0.134***	0.192***	0.183***	0.187***
	(0.003)	(0.008)	(0.012)	(0.005)	(0.007)	(0.009)	(0.003)	(0.006)	(0.003)
School-average achievement × NC Pre-K funding	−0.008*	0.016***	0.001	−0.011***	0.008	0.002	−0.007*	−0.003	−0.008*
	(0.003)	(0.005)	(0.006)	(0.003)	(0.006)	(0.005)	(0.003)	(0.005)	(0.003)
Observations	1,149,113	331,257	81,268	669,376	67,211	270,689	878,424	93,492	1,055,621
Federal PPE									
NC Pre-K funding ($000s)	0.013	0.020	0.039**	−0.003	0.003	0.025^{+}	0.006	0.009	0.014
	(0.011)	(0.021)	(0.014)	(0.009)	(0.020)	(0.014)	(0.011)	(0.018)	(0.010)
Federal PPE (in 2019 $000s)	−0.128***	−0.118***	−0.038	−0.149***	−0.161***	−0.106***	−0.129***	−0.163***	−0.124***
	(0.022)	(0.023)	(0.035)	(0.023)	(0.032)	(0.018)	(0.024)	(0.022)	(0.023)
Federal PPE × NC Pre-K funding	0.032*	0.020	0.014	0.057***	0.009	0.039*	0.029*	0.034*	0.032*
	(0.014)	(0.016)	(0.025)	(0.015)	(0.013)	(0.019)	(0.013)	(0.017)	(0.014)
Observations	1,149,113	331,257	81,268	669,376	67,211	270,689	878,424	93,492	1,055,621
Teachers with <3 years experience (in 10%)									
NC Pre-K funding ($000s)	0.027*	0.030*	0.036**	0.015	0.008	0.042***	0.019^{+}	0.021	0.027**
	(0.010)	(0.014)	(0.014)	(0.009)	(0.020)	(0.012)	(0.011)	(0.017)	(0.010)
Teachers with <3 years experience (in 10%)	−0.052***	−0.046***	−0.048***	−0.048***	−0.061***	−0.038***	−0.052***	−0.054***	−0.051***
	(0.004)	(0.005)	(0.009)	(0.004)	(0.007)	(0.004)	(0.004)	(0.006)	(0.004)

(*Continued*)

TABLE 17. (*Continued*)

	Race/Ethnicity					Mother Education		Birthweight	
	Full Sample (1)	Black (2)	Hispanic (3)	White (4)	Other Races (5)	Less Than High School (6)	High School or More (7)	Low (8)	Normal or High (9)
Teachers with <3 years experience × NC Pre-K funding	0.008*	0.002	0.005	0.004	0.006	0.003	0.009*	0.010*	0.008*
	(0.003)	(0.004)	(0.005)	(0.004)	(0.005)	(0.004)	(0.004)	(0.005)	(0.003)
Observations	1,187,543	344,099	81,574	693,627	68,242	279,037	908,506	96,126	1,091,417
Percent of SNAP recipients									
NC Pre-K funding ($000s)	0.001	−0.001	0.020	−0.022	0.003	0.005	−0.008	−0.013	0.002
	(0.019)	(0.039)	(0.021)	(0.018)	(0.029)	(0.025)	(0.018)	(0.027)	(0.019)
% of SNAP recipients in the year of the test	−0.022***	−0.022***	−0.018**	−0.025***	−0.021***	−0.028***	−0.020***	−0.030***	−0.021***
	(0.005)	(0.004)	(0.006)	(0.005)	(0.005)	(0.005)	(0.005)	(0.004)	(0.005)
% of SNAP recipients × NC Pre-K funding	0.004*	0.003+	0.004	0.006***	0.002	0.005*	0.004*	0.005*	0.004*
	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
Observations	1,207,336	349,271	82,485	706,534	69,045	283,268	924,068	97,599	1,109,737

Note. Standard errors are given in parentheses. Standard errors clustered at the county level. All models include Smart Start funding (in $000s), child and family covariates, county covariates, and county, program year, and quarter fixed effects. See Table 3 for details. Note that the “Other races” subgroup shown in column 5 combines children identified as either Asian, Native American or Mixed race due to small sample sizes for each respective group.
HS = Head Start; NC Pre-K = North Carolina Pre-K; PPE = per-pupil expenditure; SNAP = Supplemental Nutrition Assistance Program.
$^{+}p < .10$.
$^{*}p < .05$.
$^{**}p < .01$.
$^{***}p < .001$.

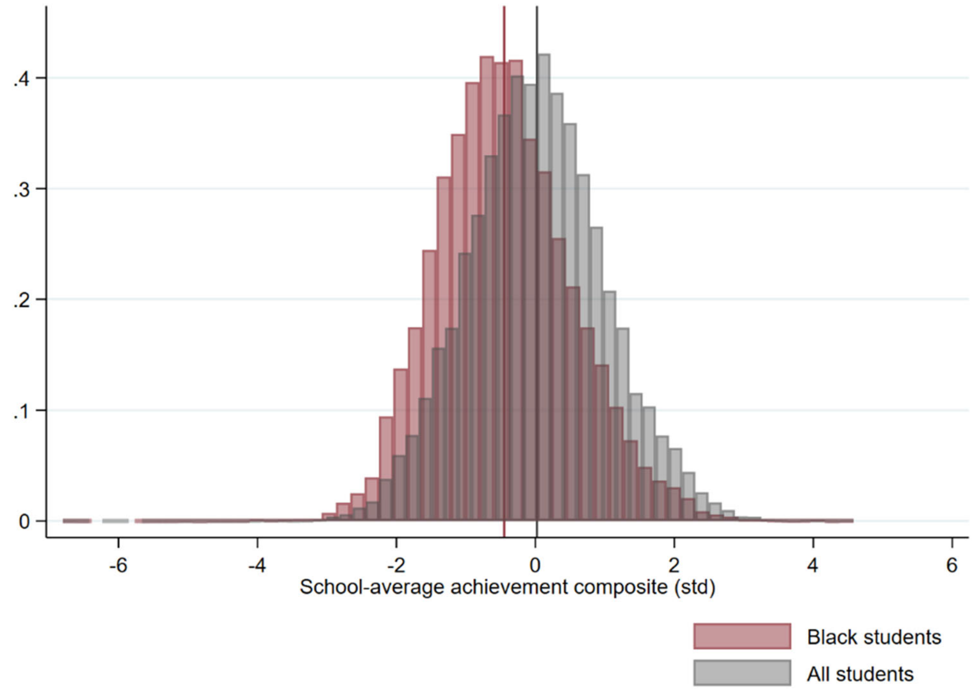


FIGURE 5.—Histogram of school average achievement composite for full sample of children and Black children.

average achievement (i.e., >1 *SD*). For these students, NC Pre-K funding produced a negative (but nonsignificant) effect on achievement, similar to the compensatory effect implied by our interaction results estimated for the full sample. Instead, the apparent positive interaction between pre-k funding and school-level achievement seems to have been driven by moving from the lowest performing schools to the middle of the distribution. For Black students attending schools in the low end of the school achievement distribution, the coefficient on NC Pre-K funding was positive, but smaller and not significant (.02). For Black students attending schools in the middle of the performance distribution, the funding effect was larger and significant (.042, p <.05), indicating that NC Pre-K effects were most robust for Black students attending middle range schools. In summary, there appears to be complementarity between pre-k funding and school average achievement for Black students moving from low- to middle-achieving schools, but the effect reverses when students move from middle- to high-performing schools.

Summary of Environmental Heterogeneity Analyses

In this chapter, we tested our hypotheses regarding complementary, compensatory, and additive effects across a host of environmental moderators.

For nine moderator variables, we found evidence for additive effects; that is, hypothesized moderator variables produced significant main effects on student achievement outcomes but nonsignificant interactions with NC Pre-K funding. We also found evidence to support a compensatory effect for four moderator variables. Indeed, NC Pre-K funding protected children from later adverse effects of low school achievement, low school-level federal funding, lower levels of teacher experience, and higher rate of SNAP benefits in the county. In other words, larger effects of NC Pre-K were observed under adverse conditions. Findings were generally similar for math and reading achievement, though effects were larger for math in several cases. In the final chapter, we review these findings and place them into the broader discussion about scale-up of public pre-k programs.

VII. Considering the Implications of the NC Pre-K Findings for Developmental Theory and Continued ECE Program Scale-Up

This study is motivated by the mystery regarding the role of environmental factors in shaping the longer-run impacts of scaled-up public prekindergarten program exposure. Some prior studies had shown stable positive effects on children's academic outcomes, some had shown fadeout of these effects within several years after pre-k, and some had shown positive effects that emerge later in life (Phillips & Pre-Kindergarten Task Force, 2017). After reviewing the literature, we concluded that all of these effects may be valid, but only under the conditions in which those studies were conducted. We hypothesized that the effects of pre-k may vary across subgroups of children and as a function of environmental factors that are present before, during, or after the pre-k year. Our study leveraged variation in the scaling up of North Carolina's pre-kindergarten program across two decades as a natural experiment to test the association between the state's pre-k funding level in a given county and year and the elementary-school academic outcomes of the children residing in that county and year. We examined the achievement records of all public-school children in NC's 100 counties across 18 years, and used regression models with fixed effects to test associations.

Hypotheses from prior work suggest several possible directions of inquiry. One theory, which we term "dynamic complementarity," posits that two environments that exert positive influences on a child's development will interact such that the positive effect of one environment will be further enhanced by the positive effect of the other environment (i.e., positive interaction), thereby generating a multiplicative (synergistic) benefit. Such complementarity would occur if the benefits of exposure to NC Pre-K were further enhanced by exposure to a particularly high-quality schooling environment during elementary school. Complementarity effects may also operate if we observed that children from homes that provided more opportunities for positive early cognitive and behavioral development benefitted the most from pre-k exposure.

A second theory poses that pre-k may act to compensate against the potential negative effects of environments that children experience before, during, or after pre-kindergarten. This idea is related to the theory of "substitution" described by Bailey, Duncan, and colleagues (2020). The psychological theory of buffering suggests pre-k may act in a compensatory way for children who previously experienced an adverse environment. The psychological theory of protective factors suggests pre-k may protect children from future adverse conditions they may encounter. This theory of compensatory interactions also predicts that pre-k exposure provides fewer marginal benefits if a child has access to other environments that provide similarly enriching experiences. In effect, a high-quality pre-k experience would have the largest benefits for children who otherwise would encounter the lowest quality educational environments. This theory of compensatory effects asserts that pre-k exposure buffers children from the adverse effects of prior experiences (i.e., pre-k minimizes the prior adverse effect) and prophylactically protects the child from future adverse experiences (i.e., pre-k minimizes the harm done by future adverse experiences).

A third theory poses that pre-kindergarten exposure has a positive impact on a child's outcomes that is independent of prior, concurrent, and future significant experiences; this circumstance would be represented as an "additive" effect to pre-k no matter what other experiences a child has.

What We Found

Our findings are summarized in Table 18. Here, we present interaction findings for each of our key classes of moderators. We note whether each moderator produced a statistically significant main effect and the direction of the effect in models predicting 5th-grade achievement. Importantly, we note whether each moderator produced a statistically significant interaction with our NC Pre-K funding variable, the direction of the interaction coefficient, and whether we interpret the interaction to be consistent with one of the key three hypotheses described above (i.e., additive, compensatory [buffering, protective], or complementary effects). For interactions with child and demographic characteristics (e.g., low birthweight, mother's education, race/ethnicity, sex), we note the pattern in the interpretation column but provide more details regarding our interpretation of these effects below. Because observing a significant main effect for a proposed moderator was essential for understanding the theoretical meaning of any interaction with NC Pre-K funding, we simply state "no main effect" for variables that did not produce a statistically significant effect on achievement.

A Significant Main Effect of Exposure to NC Pre-K Funding for Achievement With Null Effects for Grade Retention and Special Education Placement

We found that funding for NC Pre-K significantly predicted higher academic achievement test scores in 5th grade, and the overall magnitude of association was large enough to matter for policy and practice. Furthermore, this association held for most of the groups considered (Black and Hispanic students; boys and girls; normal/high birthweight; children with low mother's education). We did not find a significant effect of pre-k funding for children in the "other races" group or children with low birthweight. The effects for

TABLE 18

Summary of Key Findings for Moderation Effects of NC Pre-K Funding on Achievement

Variable	Main Effect of Moderator Variable	Direction Of Main Effect	Interaction w/NC Pre-K Funding	Direction	Interpretation of NC Pre-K Effect
Child characteristic					
Black (1 = yes, 0 = no)	Yes	–	Yes	+	Pre-K effect stronger for Black children
Hispanic (1 = yes, 0 = no)	Yes	–	Yes	+	Pre-K effect stronger for Hispanic children
Male (1 = yes, 0 = no)	Yes	–	Yes	–	Pre-K effect stronger for girls
Mother with no high school diploma (1 = no diploma, 0 = diploma)	Yes	–	Yes	+	Compensatory
Low birthweight (1 = yes, 0 = no)	Yes	–	No[a]		Additive[a]
Early childhood education exposure					
Smart Start funding	No		No		No main effect
Early Head Start exposure	Yes	–	No		Additive
Head Start exposure	No		No		No main effect
School environment					
School-level achievement	Yes	+	Yes	–	Compensatory
Federal funding	Yes	–	Yes	+	Compensatory
State funding	Yes	–	No		Additive
Local funding	Yes	+	No		Additive
Teacher environment					
Teacher–student ratio	Yes	+	No		Additive
% National Board certified	Yes	+	No		Additive
% <3 years experience	Yes	–	Yes	+	Compensatory
% teacher turnover after 1 year	Yes	–	No		Additive

(Continued)

TABLE 18. (*Continued*)

Variable	Main Effect of Moderator Variable	Direction Of Main Effect	Interaction w/NC Pre-K Funding	Direction	Interpretation of NC Pre-K Effect
Economic environment					
Job loss	No		Yes*[a]	+	No main effect
Median family income	Yes	+	No		Additive
% SNAP	Yes	−	Yes	+	Compensatory
% Medicaid	Yes	−	No		Additive

Note. The main effect column denotes whether the main effect was statistically significant for a given moderator. The direction column denotes whether the main effect positively (+) or negatively (−) predicted Grade 5 achievement. The "interaction w/NC Pre-K funding" column denotes whether the interaction was statistically significant at the .05 level in our Chapter IV analyses. For child demographic characteristics, the interaction column came from Table 6. The "Direction" column denotes whether the interaction produced a positive (+) or negative (−) prediction to Grade 5 achievement. Finally, the "interpretation of NC Pre-K effect" column denotes whether we view the interaction as evidence consistent with the complementarity, buffering, or an additive model.
NC Pre-K = North Carolina Pre-K; SNAP = Supplemental Nutrition Assistance Program.
[a]Instances of inconsistent evidence. For low birthweight, the subgroup model suggested larger effects for the "normal/high" birthweight group, while the interaction model suggested larger pre-k effects for the low birthweight group. For job loss, the interaction was only observed for one of the two tested interaction models.

White students and students whose mothers had a high school education were positive and at the margins of statistical significance.

To examine whether NC Pre-K enrollment was associated with achievement gains, we supplemented our main analyses by estimating an IVs analysis that used variation in funding as an instrument for enrollment. We found that an additional $1000 of funding for NC Pre-K raised the probability of enrollment in the program by approximately 15 percentage points, on average, and this effect on enrollment held for every demographic group considered. Next, we found that enrollment linked to increases in funding was associated with a gain of approximately 0.20 *SD*s in academic achievement in 5th grade, and this enrollment effect was shared for almost every demographic group within our sample (mirroring the subgroup results for the main funding measure described above). However, these IVs effects should be interpreted with some caution because the "exclusion restriction" may not be fully satisfied in our study, given the possibility that funding also affected student achievement through spillover effects. We view these enrollment findings as projections for the effect of enrollment on student achievement that warrant further replication from other rigorous designs that can identify the causal effects of enrollment (e.g., RCTs). In sum, our models of student achievement indicate that additional funding was associated with increases in enrollment in the NC Pre-K program, and the funding boost to enrollment was associated higher academic achievement 6 years later.

Contrary to expectations, we found no detectable effects of funding on measures of grade retention and special education in 5th grade. Odds ratios for both outcomes were near "1," indicating that in our updated models with additional later cohorts, funding increases did not lead to differences in placement rates for these two crucial educational tracks.

Additive Effects of NC Pre-K Funding Across Most Tested Moderators

Beyond updating the main effect estimates for NC Pre-K, our other primary objective was to test whether NC Pre-K funding effects were enhanced or diminished when children encountered a range of other environmental experiences outside of pre-k. Across most of the moderator tests we pursued, we found evidence that NC Pre-K effects were additive when considered alongside other environmental exposures. In other words, NC Pre-K effects were discrete and not significantly moderated by other environmental experiences considered here. Below, we review the key moderation tests that produced evidence for "additive effects."

Access to Alternative Forms of ECE

Beginning with our analyses that examined access to alternative ECE programs at the county level, we found no significant interactions between NC Pre-K exposure and our measures of Smart Start funding, EHS enrollment, and HS enrollment. We also found no indication that NC Pre-K funding effects were enhanced or diminished in counties that had larger HS footprints when the program initially rolled out. The lack of moderation in these analyses was perhaps due to the requirement that NC Pre-K funding was initially provided with the condition that administrators partner with local HS and Smart Start initiatives to reduce redundancy (Peisner-Feinberg, 2003). To this end, funding for NC Pre-K slots often flowed directly to local HS centers to supplement their reach (Peisner-Feinberg et al., 2020). Thus, the lack of interaction effects suggests that the NC Pre-K program produced consistent achievement effects regardless of a county's level of participation in these other ECE programs. One might have expected that funding for the programs could be redundant, as HS, Smart Start, and NC Pre-K share many of the same goals and primarily target the same children, and so could have produced diminishing effects when combined in the same environment. Certainly we saw no sign of "substitution" in this set of analyses.

School and Teacher Quality

Across many of the elementary school measures considered, we also detected no evidence for interactions with NC Pre-K funding. Indeed, state and local funding for elementary schools, student–teacher ratio, having more Nationally Board certified teachers, and teacher turnover rate all produced main effects on children's achievement, but none of these measures produced statistically significant interactions with funding for pre-k. These findings suggest that across levels of school and teacher quality, NC Pre-K effects on achievement were consistent. Thus, children benefitted from more exposure to NC Pre-K regardless of whether their elementary school had more state funding or whether their school had lower teacher turnover rates. It should be noted that some measures of school quality *did* produce interactions that we review below, making the findings for our school quality analyses more equivocal when compared with the analyses for alternative forms of ECE.

Exposure to Community Economic Conditions

For median family income and proportion of families on Medicaid, we again detected significant main effects of these variables. Not surprisingly, a high median family income in a community was associated with higher academic achievement scores. Further, living in a community with a high proportion of families receiving Medicaid was associated with lower academic achievement scores. It should be noted that we do not interpret this effect as an indication that participation in Medicaid lowers children's achievement. Rather, increased Medicaid participation in one's county is largely a function of the level of wealth and income in a county, and lack of general economic resources in a child's environment should be expected to correlate negatively with children's achievement for a variety of reasons (e.g., poorer schools, less opportunities for learning outside of school, etc.). In both of these cases, we observed no significant interactions with NC Pre-K funding, thus supporting a model of additive effects.

Significant Interactions Tend to Provide Evidence for the "Compensatory" Hypothesis

Four of our analyses supported a model of compensatory effects of pre-k, indicating that funding for NC Pre-K provided the largest benefits to children who were otherwise likely to face more disadvantage in other environments. As we explained in Chapter III, these effects could have also arisen if we had found that children from NC Pre-K exposure (i.e., the inverse).

Demographic and Family Characteristics

We begin by noting interactions with several key demographic characteristics. It is crucial to underscore that we interpret our variables indicating race as proxy measures of larger systemic forces that children are likely to encounter in the United States due to their racial profile. Thus, we interpret the differential effects for Hispanic and Black students as compensatory effects, not because we view these racial labels as synonymous with "disadvantage," but rather because we understand that children from these groups are likely to face more environmental disadvantage due to factors such as structural racism and economic inequality. With this caveat in mind, we found that the two groups who experienced the largest benefit of enrollment in NC Pre-K were students from Hispanic families and students whose mothers had not graduated from high school. This evidence suggests a compensatory effect for NC Pre-K, whereby exposure to NC Pre-K has larger benefits for children who were likely to grow up in more impoverished economic environments or face inequities in opportunity resulting from structural racism. Moreover, the NC Pre-K interaction coefficient for children identifying as Black was also statistically significant and positive (Table 6), though it was not as large in magnitude as that found for Hispanic students.

The larger effects for Hispanic students may also further reflect a benefit of the program for DLL. Indeed, the recent NC Pre-K RCT study also found that the largest early language and literacy gains were enjoyed by children who were DLL (Peisner-Feinberg et al., 2019). Unfortunately, we did not have a direct measure of DLL status in pre-k in our current study. However, future work should continue to examine this issue because children who are less likely to receive exposure to English language in their home environment may benefit more from English or Bilingual instruction in pre-k.

In terms of other child characteristics pulled from the birth records, we found equivocal results by birthweight. Across our models, this was the lone finding that did not produce consistent results across the two approaches for testing moderation (i.e., splitting the sample vs. testing the interaction; see footnote #6), and given the small magnitude of the difference, we think it is most prudent to disregard apparent differences based on birthweight.

Elementary School Average Performance

Although we found that exposure to pre-k funding had significant benefits for children's 5th grade achievement, we found these positive effects were diminished when children attended higher performing elementary schools (i.e., reduced marginal benefit). This finding also indicates that the effect of exposure to high levels of NC Pre-K funding protects a child from the adverse effect of later attending a low-achieving school. The pattern suggests that elementary school experiences may act as a partial substitute for early childhood educational experiences (see Bailey, Duncan, et al., 2020). Because the interaction also suggests that pre-k funding blunted the potential negative effects of attending a lower-performing elementary school, this finding is further evidence for a compensatory effects model of pre-k.

Importantly, this compensatory finding may carry implications for understanding the mechanisms of pre-k effect fadeout often characterized in the ECE literature (Bailey et al., 2017). Our negative interaction effect between school achievement and pre-k funding exposure suggests that children who did not attend pre-k but entered high-performing schools likely “caught up” to pre-k attenders over the course of elementary school. Unfortunately, we lack the consistent measures of achievement from pre-k through Grade 3 to accurately characterize how this pattern of catch up might have worked in higher performing schools, but this finding seems to cast doubt on the idea that policies promoting elementary school academic performance will help sustain the positive effects of pre-k. Instead, children who attend higher performing schools may have relatively less to gain academically from pre-k exposure in terms of building academic skills, but children likely to attend low performing elementary schools may see the most benefit from pre-k attendance. We return to this point below.

Attending a School with Inexperienced Teachers

Here we also found support for a compensatory model. Children who attended schools with a high proportion of teachers with fewer than 3 years of teaching experience had lower academic achievement scores. At the same time, the benefits of exposure to NC Pre-K funding are greater for children who attended elementary schools with a high proportion of inexperienced teachers. Thus, exposure to NC Pre-K funding protects a child from the future adverse effect of attending a school with a high proportion of less-experienced teachers.

Federal Funding

The effect of federal per-pupil funding is more complex. Most federal funding is directed to special education and other compensatory interventions for districts with high poverty and adversity. Consistent with this point, we found a negative association between federal funding levels and children's academic achievement, which is very likely a function of higher federal funding going to school districts that are in greater need. (A negative effect for state level school spending was also detected, again likely due to compensatory funding mechanisms.) Indeed, we also found that at the county-level, increased federal per-pupil funding for school was positively correlated with other markers of county-level ED, like receipt of SNAP benefits. Thus, paradoxically, higher federal funding should be considered an indicator of need rather than a driver of benefit. Consistent with our expectation of compensatory effects, we found a significant positive interaction between federal funding and NC Pre-K funding that we interpret as showing that the effect of NC Pre-K funding was larger in schools that were in higher need. This is consistent with our finding for school-level achievement, as pre-k funding helped children more when they attended elementary schools with greater need for federal funding, supporting the compensatory model of pre-k effects.

Community Economic Adversity

Also consistent with the compensatory hypothesis, we found that NC Pre-K effects were enhanced in communities that had higher percentages of residents receiving SNAP benefits. This pattern suggests that children living in communities marked by substantial ED benefitted the most from investment in the NC Pre-K program.

As we noted above, effects for job loss were somewhat mixed. In one of our tested models, we found that NC Pre-K effects were enhanced in counties that experienced more jobs losses. However, we temper our interpretation of these findings given the lack of a significant main effect for the job loss measure. Further, the interaction between job losses and NC Pre-K funding was significant only in models that included time trends as controls, further suggesting some sensitivity in the interaction finding. It should be noted that previous work by Ananat and colleagues (2011), which relied on a similar county fixed effects design, detected negative effects of job losses on student achievement in Grade 8. Thus, it is likely that community job losses have negative effects on achievement for affected students, and the positive interaction between job losses and NC Pre-K funding detected in one of our models would suggest a protective effect. However, given the inconsistencies in our estimates, this evidence should be interpreted cautiously.

Little Evidence for Dynamic Complementary Effects of Pre-K Funding

Across the interaction models tested, we found almost no support for dynamic complementarity. This may be seen as surprising because the "sustaining environments" hypothesis is ubiquitous in the recent pre-k research literature. We return to this issue below, but it should be noted that two analyses did turn up evidence that could be consistent with complementarity, however we believe the theoretical implications from each test are unclear.

In the first case, dynamic complementarity was arguably found for the interaction between pre-k and the sex of the child. Pre-k funding had positive effects for both girls and boys, but girls had better achievement outcomes than boys, and the effect of pre-k exposure was greater for girls than boys. Given the mixed literature on whether ECE experiences have stronger effects on girls than boys, we have no evidence-based theory to explain this pattern. Further, given the complicated nature of gender dynamics in relation to academic achievement, we do not believe this finding constitutes strong evidence for complementarity.

Second, for each of the four environmental moderators that produced statistically significant interactions (i.e., school achievement, federal PPE, % teachers with less than 3 years of experience, and % SNAP recipients) we also tested whether these interactions were consistent for each of the key demographic subgroups featured in Chapter V. Across most of the subgroups, we found that the tested environmental moderator produced an effect that was consistent with the finding for the larger sample. However, for Black students, we found that school achievement produced a *positive* interaction with NC Pre-K funding.

Although this interaction suggested that for Black students, NC Pre-K effects on achievement were enhanced in higher quality elementary schools, post-hoc descriptive analyses revealed that Black students attended lower-quality elementary schools, on average, than the full population of students, and this quality difference amounted to approximately 0.70 *SD*s on the school achievement measure when comparing Black to non-Black students. Shown in Table A21, the largest NC Pre-K effect appeared to exist for Black students attending schools in the middle of the pack in terms of performance. However, for Black students who were in "high performing" schools, the effect of NC Pre-K was *negative* (though not significant). Thus, we would characterize this evidence as potentially exposing the limits of compensatory effects, but it falls short of being "proof" of dynamic complementarity. It rather appears that for Black students, the effect of NC Pre-K was somewhat blunted (though still positive) in the lowest-performing schools. This is concerning, and suggests that the compensatory nature of pre-k may have limits for some students who face especially high levels of disadvantage in other environments. These findings certainly warrant more consideration in future work.

Similar Results for Mathematics and Reading

Before moving on, it should be noted that we only featured results for the "achievement composite" measure in Chapter VI, but we did replicate all of our key analyses for separate measures of 5th grade mathematics and reading in the Appendix. By in large, we found that results were consistent across both achievement domains. Main effects of NC Pre-K were statistically significant for math and reading, though effects for reading were slightly larger. For our moderation tests, we found that results were consistent for math and reading for three of the four key moderators that produced significant main effects and interactions on the composite variable of achievement (i.e., school level achievement, federal PPE, percentage receiving SNAP benefits). The only difference was found for the teacher experience moderator, as this interaction was only statistically significant for math achievement. Likewise, the job loss interaction was only detected for math achievement, but this interaction was, again, not consistently observed across various specifications using our composite achievement measure, and failed to produce a main effect for either math or reading. In general, we did not see a clear pattern of distinction between results for math and reading that would warrant us drawing unique conclusions for the two domains of achievement.

Interpreting What We Found

Across most models tested, the findings largely supported an additive model that asserts that funding for scaled-up, statewide pre-k in North Carolina provides a boost in longer-run academic achievement for most groups of children under most past and future negative environmental circumstances. Because the additive boost is especially large for children from disadvantaged backgrounds or living in impoverished circumstances, higher funding is also likely to reduce disparities across groups. For children who experienced contexts of disadvantage, this funding partially compensated for the loss they would have otherwise expected. The findings, in general, did not support the hypothesis that pre-kindergarten funding would catalyze, synergize, or dynamically complement other resources. We found virtually no support for the sustaining environments hypothesis. We interpret these findings in terms of three models of investment in children's educational development.

NC Pre-K As an Additive Investment

The findings support an additive model of the role of investments in children's learning that is consistent with production function models, investment models, and environmental support models. These models are blunt: inputs contribute to outputs in a main-effect, additive way. An additive model asserts that each input supplements children's development regardless of other circumstances or inputs. The majority of findings in this study are consistent with this additive model, as the effect of NC Pre-K is uniform and unaltered across levels of other environmental experiences considered.

Regardless of interactions with other investments, the finding of positive effects of pre-kindergarten exposure on academic achievement at the end of elementary school is remarkable in and of itself and may be attributed partially to the fact that the first 5 years of life are a sensitive period in development when children are particularly receptive to learning new skills that endure across the lifespan and guide future learning (Brown & Jernigan, 2012). Environmental input has an inordinate lifelong impact during these early years (Huttenlocher et al., 1998; Magnuson & Waldfogel, 2008), especially in the presence of early disparities by SES in children's skills prior to school entry (Garcia & Weiss, 2017; Hanushek et al., 2019).

These findings are consistent with results from selected studies reporting the effects of good-quality, scaled-up, pre-k programs (e.g., Cascio and Schanzenbach, 2013; Gormley et al., 2018), though they contrast with other results from recent RCTs of good-quality, scaled-up pre-k programs (e.g., Lipsey et al., 2018; Weiland et al., 2020). The positive effects on elementary-school achievement could be attributed to the program's high quality features in NC. Individual features of the program could not be tested here, but characteristics of pre-kindergarten programs that could be responsible for enduring impacts include strong accountability that accrues from state oversight, alignment with the K-12 educational system that comes from the design of the program by officials who also oversee elementary school curricula (Kauerz, 2008), credentialed teachers who spend time engaging children in cognitively stimulating activities with a known curriculum (Phillips et al., 2009), single-age classrooms that are not diluted by younger peers who depreciate teaching levels (Ansari et al., 2016), and high overall quality as rated by classroom observers (Xue et al., 2016).

Questions of how to measure and classify preschool quality has been central to research in the field. NIEER has laid out guidance regarding structural (e.g., group size, teacher–child ratio, teacher credentials) and process (e.g., curriculum, children's positive and stimulating interactions with caregivers) features that define high-quality state pre-k programs. Structural qualities on their own have not been associated with lasting and sizable impacts on student learning (Early et al., 2007; Pianta et al., 2005; Sabol et al., 2013); whereas features of process quality, particularly teacher–child interactions that are warm, responsive, and support learning, have shown stronger positive correlations with children's learning outcomes (Early et al., 2007; Mashburn et al., 2008; Yoshikawa et al., 2013).

Those intriguing findings aside, our study analyses cannot reveal why program impacts were strong and consistent, providing instead a classic "black box" evaluation of these impacts.

NC Pre-K as Compensation (Buffer and Protection) Against Other Environmental Disadvantage

Although the positive impact of pre-k funding holds in a general way, we find important conditions under which the impact is especially strong, consistent with the compensatory model. To this end, we found that funding for NC Pre-K had an inordinate positive impact on children who were born into circumstances of financial disadvantage or historically-based cultural discrimination (e.g., being Black or Hispanic, or born to a mother without a high school degree). These findings are in line with many other evaluations of pre-k effectiveness (see review in Duncan et al., 2022), and they provide further evidence that pre-k largely acts as a compensatory intervention for children from disadvantaged backgrounds. This is not news, and our findings certainly reflect the compensatory goals of NC's program.

However, our findings regarding children's subsequent environmental experiences have further sharpened our understanding of the compensatory function of pre-k. Indeed, contrary to recent expectations in the field regarding dynamic complementarity and sustaining environments, we found that exposure to NC Pre-K funding partially protected children from the ill effects of *future* adversity. The effect of NC Pre-K funding was stronger among children who later attend an elementary school that typically had poor academic performance or a school that received high levels of federal funding (which is usually allocated to districts in highest need). Pre-K funding was also stronger for children who lived in communities with more ED during elementary school, as measured by the percentage of families receiving SNAP benefits (with some indication that communities with more job losses also benefitted most).

Few studies, to our knowledge, have ever examined whether protective interactions between ECE and markers of environmental disadvantage extend to environments measured *after* children leave preschool. Indeed, our results suggest that access to high-quality ECE may partially protect children against environmental disadvantage, regardless of whether that disadvantage is encountered before or after the preschool experience. As we suggested in Chapter III, these findings suggesting protective effects should not be entirely surprising, given that many studies have found that pre-k effects are often largest for children from demographic groups marked by environmental disadvantage. Although we often treat demographic characteristics as measures that capture experiences of children before pre-k begins (i.e., "baseline" characteristics), there is little reason to think that the environmental influences of racial or economic inequality stop after pre-k ends.

Of course, the reciprocal assertion of a buffering model is that the impact of investments attenuates, or asymptotes, under conditions of initial comparative advantage. In this study, those conditions include being born into a White family of (presumed) cultural privilege, being born to a mother with a high school diploma, and attending a high-achieving elementary school. Thus, we found that pre-k does not support a "rich get richer" model of development. Below, we further consider where our findings, which were most consistent with the additive and compensatory models described above, depart from prevalent thinking in the field regarding dynamic complementarity.

Nonsupport for a Model of Dynamic Complementarity for Pre-K

Outside of the aforementioned school achievement effect observed for Black students attending very low performing schools (<-1 *SD* statewide school average achievement), we did not find much evidence to support a dynamic complementarity model, in spite of the strong theoretical basis for such a model. Indeed, rigorous evidence of dynamic complementarity has been reported from other studies with sound causal designs (Johnson &

Jackson, 2019), yet we found little indication that benefits of NC Pre-K were enhanced in the presence of other high-quality environmental exposures after preschool.

However, we do not take this to mean that dynamic complementarity should never be expected when considering the sequence of developmental opportunities. Rather, we have narrowed the circumstances under which we might predict complementary interactions between two environmental settings. As we discussed in Chapter III, complementarity may be the more likely model in the case of hierarchical skill-building, when a child encounters two successive environmental settings that provide highly sequenced curricular input (e.g., Clements et al., 2013). However, it appears that such circumstances are rarely measured in educational studies, and it is not clear that such settings even exist in most schools outside the bounds of carefully executed curricular interventions (see Engel et al., 2012). Thus, it remains possible that dynamic complementarity could play a role if formative assessment guided teachers' instruction, and children encountered content that matched the developmental sequence of learning trajectories, as proposed by Bailey et al. (2020) and Clements et al. (2013).

Statewide pre-kindergarten initiatives do not typically bring the specific, highly-sequenced curricula that would be necessary to observe dynamic complementarity. Rather, they bring a safe and structured care environment, which enhances social-emotional functioning or relieves stress in the home, as well as opportunities for learning. From this perspective, pre-k may provide a more general environmental intervention for children. Only a handful of studies searching for evidence of sustaining environments have considered moderating factors that directly measure instruction or curricular content (Burchinal et al., 2022). Here, we considered measures of the general school environment, economic conditions of the community, and availability of alternative forms of ECE. Thus, our findings match the conclusions of a recent review of "sustaining environments" evidence in ECE studies: General measures of subsequent environmental quality typically yield nonsignificant interactions with measures of ECE exposure (Bailey et al., 2020). Even so, we leave open the possibility of dynamic complementarity if pre-k and elementary school curricula are closely aligned, and we suggest that state policymakers implement closer alignment to test this possibility.

Qualifications and Limitations

As with any study, our work contains important limitations that qualify the interpretation of our results. Below, we note several of the most important limitations, and discuss how they may affect implications that we discuss here.

Design Limitations

Our method of analysis rests on the assumption that by controlling for differences between time periods and counties (as well as a host of child demographic factors and changing county-level characteristics), we can isolate the causal effect of "haphazard" and plausibly exogenous differences in NC Pre-K funding over time on child achievement at Grade 5. While we suggest that this design is likely more rigorous than many other studies that have tried to use measures of child characteristics to control for selection into preschool, the study could still be susceptible to omitted variables bias from factors that are correlated with within-county changes in NC Pre-K funding over time. It is difficult to speculate as to what those factors could be, as we found little change in our key coefficient when controls for Smart Start, HS and EHS were introduced. However, unlike an RCT, our study does not fully eliminate such possibilities of bias. The "two-way fixed effects" approach has also been criticized by recent methodological papers in econometrics for estimation issues that can lead to bias (see discussion in Chapter 9 of Cunningham, 2021). Yet, it is unclear how these recent concerns apply to settings in which the independent variable of interest is continuous, which would apply to our use of "dollars" as the key measures of preschool exposure (Callaway et al., 2021). It should also be noted that any weaknesses in the identification of our funding effect are extended to the results from our IVs analysis. Because we cannot fully rule out the possibility that funding effects on achievement were also partially driven by spillovers, the exclusion restriction criteria for the IVs analysis may not be fully satisfied in our setting. Consequently, the results from the IVs analysis should also be seen as suggestive in terms of enrollment effects.

With these limitations in mind, how should we approach the causal implications of these results? In our view, causal inference operates on a continuum. Every study, including RCTs, carry concerns about internal validity, and researchers must consider those concerns against the strengths of the design. In the pre-k literature, most studies suffer from the most severe form of selection bias, where pre-k attendance and nonattendance are determined completely by participant choice, and researchers can only use regression techniques (i.e., control variables) to try and balance out the characteristics of these two groups. This approach will almost always fall short because even "good" control variables are unlikely to capture all of the unobserved differences between families who make divergent choices regarding pre-k enrollment. Although our approach also depends somewhat on control variables, our main estimates are not drawn from nonrandom comparisons between pre-k attenders and nonattenders. Rather, we examine differences in the availability of funding within counties over time. Omitted variables bias in our setting would exist if nonrandom differences at the county-level cause both fluctuations in the availability of NC Pre-K funds within a county and later differences in achievement at the child level. As we described in Chapter

V, we cannot fully rule out this possibility, and the sensitivity tests suggested that our estimates depended in part on controlling for time-varying county characteristics. However, we believe this design is likely to carry *less* bias than the typical approach of comparing pre-k attenders and nonattenders, and provides strong evidence for the broader literature by presenting findings from a design that is qualitatively distinct from what is commonly found in pre-k studies. Nevertheless, given that we cannot fully account for the exogeneity of pre-k funding, caution is warranted when deriving causal interpretations of our results.

Endogeneity of Moderators

Although our design does attempt to examine the causal effect of NC Pre-K funding on children's achievement (though we have noted several limitations to this pursuit above), we can make no such claims for the moderators considered here. In the case of most of our moderator measures, it is unlikely that the county and year fixed effects approach can identify the causal effect of a given moderator on child achievement. Nonetheless, moderators that are obviously endogenous to other factors that cause changes in children's achievement (e.g., school level achievement), still provide theoretical utility in that they capture broader variation in the quality of early childhood environmental experiences (i.e., children in higher-performing schools are likely to have access to other environments that enrich their education outside of school).

Despite these limitations, we suggest that our approach may still be an improvement over most studies that have examined how subsequent environments interact with exposure to ECE, as most studies in this area employ conventional correlational approaches with control variables (i.e., preschool status is not randomly assigned; Bailey et al., 2020). Moreover, we found that most of the moderators considered had modest associations with NC Pre-K funding when using our county and year fixed effects approach. In the case of job losses, we even relied on earlier research that had established the exogeneity of county-level job losses to child achievement (Ananat et al., 2011). Still, our moderator effects should be seen as associational evidence, and we would caution anyone against making straightforward policy interpretations based on our moderation results. For example, our analyses of teacher characteristics do not necessarily imply that schools should increase the number of Board Certified Teachers as a mechanism for improving student achievement, despite the positive and statistically significant coefficient for the Board Certified Teachers measure.

To this end, our understanding in this area would be improved by studies that can isolate exogenous variation in both the initial exposure to ECE and the subsequent environment. Yet, we recognize that such designs are rare and difficult to execute. Perhaps the best design for this line of inquiry would involve a 2×2 randomized experiment, by which access to a given early childhood program would be completely randomly assigned against exposure to a hypothesized follow-through program. This design would identify the main effects of both the ECE program and the follow-through program, but it would also identify the effects of an exogenous interaction between the two programs. Several randomized evaluations of the Building Blocks early math program have used similar designs to test the efficacy of various "follow-through" interventions that build on the Building Blocks preschool curriculum (Clements et al., 2013; Mattera et al., 2021). Unfortunately, these RCTs did not employ a full 2×2 design, preventing one from observing the effects of a given follow-through treatment *without* the pairing of the early childhood treatment (i.e., follow-through was only given to students who already had access to the preschool program). This makes it difficult to know if positive effects for the follow-through condition were driven by the unique combination of both an early childhood program and a follow-through program, or if access to the follow-through program alone may have produced additional positive impacts.

We know of one natural experiment that examined an arguably exogenous interaction between access to an early childhood program (HS) and subsequent high-quality educational environments (increases in K-12 school spending) (Johnson & Jackson, 2019). This study leveraged policy-induced changes in both HS and K-12 spending and demonstrated that the effects of both programs were enhanced when offered in combination. To our knowledge, this study likely represents the best evidence for dynamic complementarity given the rigorous identification approach, and we hope more studies can capitalize on similar designs in the future.

Outcome Variables and Age

Our study is limited by reliance on a limited set of outcome variables: an academic composite variable (average of reading and math standardized achievement test scores), special education placement, and grade retention in Grade 5. Because we found null main effects of the program on special education and grade retention, all of the moderator analyses in Chapter VI were pursued only on the achievement composite measure. We recognize that the benefits of pre-k may be much broader than what might be captured by our set of outcomes derived from administrative data. Indeed, measures of specific academic skills, social-emotional functioning, and observations of social behavior all deserve attention when considering pre-k effectiveness. Moreover, our study is limited by the sole focus on child outcomes; understanding how pre-k affects parents and families (e.g., maternal employment) are equally important when considering the efficacy of ECE initiatives (Burchinal et al., 2022).

From our view, the 5th grade age point is critical because performance by this age predicts later ecologically important outcomes such as high school graduation and typically marks the end of elementary school in North Carolina.

It is past the age at which most other studies have identified fadeout of impacts, and thus represents a "longer-run" effect. It remains for future studies to determine whether impacts continue at later stages of development and whether other crucial domains of child development (e.g., social-emotional functioning) are similarly impacted by funding for public pre-kindergarten. (We return to this point below.)

Extrapolation

Our study spanned nearly two decades (1993–2010) during which 4-year-old children were exposed to early childhood programs in North Carolina. Communities in this state experienced remarkable economic, political, and social shifts during this period. One such shift in funding for the NC Pre-K program enabled the current evaluation design. In more recent years, funding levels have stabilized, making evaluation of the impact of participation for future cohorts possibly less valid. Thus, we cannot know whether the funding effects estimated here would extend in future years, especially if other policies were adopted by North Carolina that might also affect early childhood development or later academic achievement. Moreover, although our estimates depend on the linear association between preschool funding and student achievement, we cannot be certain if additional increases in preschool funding would benefit children if preschool were to reach full saturation levels across the population of 4-year-olds. As we noted in Chapter V, we found no indication that funding effects tailed off because the quadratic term for funding was far from statistically significant. However, extrapolating beyond the current data would be inadvisable. In most of the relevant years of funding considered here, approximately 10–30% of 4-year-old children received funded slots in NC Pre-K. The number of funded slots could increase either by saturating the group of already-eligible-but-not-enrolled children or by expanding the eligibility criteria. It is unclear whether further increases in the number of funded children would continue to be associated with increases in population-level achievement outcomes in a linear fashion, but such questions certainly warrant future research consideration.

Recommendations for Practice and Policy

Our findings suggest practice and policy recommendations in several areas related to funding allocations and alignment between pre-k and elementary education.

North Carolina Should Continue Investing in Pre-K and Prioritize Funding for Disadvantaged Students

Across our models, we consistently found evidence that the NC Pre-K program brings positive benefits for all groups of children and is most effective in boosting the achievement of children who encounter other environments marked by disadvantage. Indeed, our findings are largely aligned with the current goals of the NC Pre-K program, as eligibility is prioritized to children from lower-income families, as well as those who have other markers of early risk. Based on our results, we expect that investment in children who are most in need should continue to provide the largest returns for the NC Pre-K program. However, it should be noted that we saw consistent and positive effects of NC Pre-K funding across most demographic groups considered. This consistent pattern suggests pre-k funding may also benefit children who have fewer markers of environmental disadvantage.

Pre-K Programs Should Optimize Both Universal Impact and Reduction of Disparities

A major policy issue today is whether pre-k programs should be universally offered or targeted to particular groups. The NC Pre-K program offers a unique blend. Although slots are targeted to economically disadvantaged students (plus a comparatively small number of special-needs students and children of military families), the program attempts to have universal impact by incentivizing pre-k providers to improve their quality so they qualify for pre-k slots. Across the years of this study the average classroom participating in the program included children who were not directly funded by the program. The state's goals included positive impact on these students as well as those targeted for funding. In addition, the students who were funded through a pre-k slot entered kindergarten classrooms that were populated by peers who had not been state-funded by a pre-k slot. The state's hope was that pre-k funding for a subgroup of students would enhance the initial average skill level in kindergarten classrooms, allowing teachers to focus instruction at a higher level and supporting spillover effects on peers.

The findings in this study indicate that higher state funding levels are associated with better academic outcomes for nearly all subgroups studied, including nearly all race and ethnic groups and groups disaggregated by maternal education levels. These findings are consistent with hypothesized spillover effects that state leaders had envisioned. Furthermore, the effects of state pre-k funding appear to be stronger for disadvantaged students, suggesting the program met two goals of universal impact while also reducing disparities in outcomes across groups based on advantage.

The policy implication is that it may be possible to reach both goals by prioritizing funding for disadvantaged students if those disadvantaged students are integrated with more advantaged peers in pre-k classrooms and subsequent school.

Pre-K Funding Should Support Slots Over Classrooms Segregated by Income

Unlike other early childhood programs, the NC Pre-K program does not create, staff, or fund pre-kindergarten classrooms; rather, it provides funding to counties in a year for a specified number of "slots" in existing classrooms, so long as those classrooms meet standards for structural and process quality. One implication of this feature is that many children in those classrooms were not directly funded by the state allocation but may have benefitted nonetheless. Peisner-Feinberg and Schaaf (2009) reported that between the 2003–2004 and 2008–2009 academic years, the average proportion of NC Pre-K funded children in a classroom ranged from 67% to 83% (with the share going up over this period). This means that a nontrivial proportion of children in NC Pre-K classrooms were not actually funded by the program, but would have certainly benefitted from the quality supports, standards enforcement (above and beyond that of child care licensing alone), teacher professional development activities, and other technical assistance provided to NC Pre-K providers. The benefits of these program supports certainly extend outside of the preschool-aged classrooms, and may even benefit teachers and classrooms for younger or older children served by the NC Pre-K provider. Thus, the reach of the program likely extends beyond the number of allocated slots.

Another implication of the slot-oriented, mixed-delivery model of NC Pre-K is that although a baseline standard of quality for NC Pre-K providers is assured by the state oversight, the settings are diverse and therefore measurement is challenging. Peisner-Feinberg and Schaaf (2009) summary of program data collected between 2003 and 2009 reported that approximately half of the slots were found in public school settings, with the other slots going to a mix of private childcare centers and HS sites, with quality steadily increasing over time as the program scaled-up. Our study was not able to evaluate the relation between any specific child's experience and that child's outcome, not only because we lack the data, but also that parental choice of pre-k setting is not randomly assigned. However, our evaluation of the population-level impact of the state policy is well-grounded in a strong natural experiment design.

Funding for Pre-K Should Not be Offset by Lower Funding for Other Education Programs

Because we found that the effects of NC Pre-K funding generally added to the associations observed between other positive environmental experiences and later achievement, we conclude that most children would stand to benefit in an additive way from positive investments across developmental periods. Certainly, some of our compensatory interactions suggest that preschool benefits will be largest in settings where children encounter lower-quality environments. Thus, funding preschool in contexts in which K-12 schooling quality is lower than average would appear to be a worthwhile investment based on the interactions observed here. However, we found no indication that investments in early childhood educational experiences should be thought of as a zero-sum game. Rather, positive investments across developmental periods will most likely yield additive benefits, and in some cases, positive investments will be greater for children who otherwise would have encountered lower-quality environments.

School Districts Should Think Strategically About Goals of Pre-K and Subsequent Elementary Education

Given that pre-k increases kindergarten readiness for its participants but does not yet reach all children, the cohort of entering kindergarteners remains diverse. One could imagine public school systems adopting different strategies, goals, and incentives that are intended to lead to different lasting impacts of interventions that students experience prior to kindergarten matriculation. A school leader might be incentivized to increase the number of students who ultimately achieve highest honors (operationalized as college-going, awards, etc.); this leader might prioritize instruction to students who enter kindergarten at the highest level, through advanced curricula that "teach to the highest level," allowing students with lower skills to remain low, or drop out. The stereotypical elite private school adopts this model, which coincides with sustaining the impact of early intervention.

In contrast, a school leader might be incentivized to reduce the number of dropouts and failing students (operationalized as the number of students performing below grade-level) as part of a goal to eliminate achievement gaps. This leader might devote extra attention to students who are in danger of falling behind or remaining behind. This strategy is one of intended fadeout, leading to a diminishment of impacts of early performance-boosting interventions for some children. These strategies are incentivized by how the federal government requires that school systems tie incentives to student achievement measures, such as standardized achievement test scores. North Carolina initiated one of the earliest such incentive systems in the 1990s, called the "ABCs," which provided a model for the Bush-era "No Child Left Behind" Act. The North Carolina Department of Public Instruction evaluates schools by the proportion of students that score at grade level on EOG tests. No extra incentive is provided for a higher mean score or a higher proportion who score above grade-level. This statewide incentive therefore likely operates to minimize the continued growth of children who got a precocious boost from NC Pre-K, by attending to the needs of children who had not benefitted from NC Pre-K. Might this be the optimal education policy design? How might schools address individual differences but provide a universal, inclusive experience? Such questions are outside the scope of our study, but are certainly central to future work in this area.

Next Steps for Research

The current study suggests several high-priority areas for future research studies.

Understanding Impacts on the Entire System

The NC Pre-K program is structured so that it provides funding to centers for available slots. This organization-level funding structure presents the possibility that the program could have impacts on schools and districts that are not easily understood by merely examining effects at the individual, or child, level. Most of the potential explanations for fadeout converge on the idea of "catch-up" by control children, such that the trajectory of learning accelerates between post pre-k and follow-up for the control group but not for the intervention group. Unless there is something serendipitous about the age targeted for the pre-k intervention (that is, it is selected during the interval just prior to accelerated learning in natural development), this "catch up" pattern could be seen as a sign of a potential spillover effect for the control group.

How might catch-up in the control group be a sign of a spillover effect from the initial intervention? If the control group learns more in a given educational setting when the pre-k intervention exists, then it could be the case that "treating" some children with pre-k may make learning conditions better for their peers, even if those peers never attended pre-k themselves. In other words, if pre-k attenders and nonattenders enroll in the same educational environments after pre-k, then the accelerated catch up in the control condition might be partly caused by the presence of "treated peers." These spillover effects could occur even if pre-k attenders and nonattenders do not directly enroll in the same later educational environments, so long as higher-order educational structures (i.e., districts or schools) are able to more efficiently redirect resources toward the children who did not attend pre-k.

These counterintuitive spillover effects could occur given that educational systems are dynamic, and they may make decisions to adapt to the changing characteristics of their student base. Indeed, kindergarten teachers may recognize that many students in their classroom have already mastered basic content due to attending pre-k, and this could lead them to focus more resources on the children who did not attend pre-k. Under this scenario, the net effect of the pre-k intervention on the total population might still be positive even if differences between intervention and control children within an environment are not detectable after kindergarten. To this end, List, Momeni, and Zenou (2020) estimate that spillover effects could improve a control child's scores on cognitive tests by 0.6 *SD*s and behavioral measures by as much as 1.2 *SD*s. They conclude that unaccounted spillover effects could lead to severe underestimation of intervention effects. Thus, we need research with designs that allow for control children to be selected from truly independent environments; the environments themselves are the unit of comparison.

Indeed, there have been many highly cited RCTs in early childhood that randomize at the school level (e.g., Clements et al., 2011; Raver et al., 2011). These studies usually evaluate curricula differences between pre-k programs, and these interventions still often produce fadeout. We suggest that future research on the effectiveness of pre-k programs should consider effects on schools or districts as a unit of analysis. Such work may find that substantial support for public preschool generates positive outcomes for a district or school, even if effects fade at the individual level. Future studies in this area would most effectively address the possibilities raised here if they could clearly disentangle subsequent school environments that have high degrees of mixing between pre-k attenders and nonattenders apart from other environments that had lower levels of peer mixing. As pre-k programs continue to grow in scale across the country, we believe that understanding the role of pre-k penetration on educational systems will constitute a crucial frontier for ECE research.

Identifying Peer Effects

The interesting possibility of spillover effects driving positive effects on higher order units, such as schools and districts, also necessitates a better understanding of peer dynamics in early childhood classrooms. Future research should focus on the ways that peers influence the long-term progression of early intervention impacts. As suggested above, it could be the case that the pre-k experience leads to an initial positive impact on a child's academic readiness at kindergarten entry, and these gains in skill development indirectly support elementary-school learning of children who did not attend pre-k. It may also be the case that pre-k attenders themselves benefit by being in a classroom with more children who also attended pre-k. This may occur if pre-k attenders have more advanced language skills, fostering an enriching early childhood environment where interactions with peers lead to further gains in language and cognitive development.

Attending classrooms with higher concentrations of pre-k attenders may also benefit pre-k children by altering the content emphasis of kindergarten or 1st grade instruction. In the previous section, we imagined a scenario in which a mixed classroom of pre-k attenders and nonattenders may favor the nonattenders, as teachers redirect resources toward the children who have not already obtained school readiness skills. This leads to increased fadeout during kindergarten or 1st grade. One could also imagine a scenario where this dynamic shifts if the concentration of pre-k attenders nears 100%. Instead of targeting content toward the nonattenders, teachers may instead recognize that most of their class has already mastered the most basic content, and shift toward more advanced content that increases learning for the pre-k attenders. (But for contradictory evidence, see: Burchinal et al., 2023; Engel et al., 2012.)

To make significant progress on understanding these dynamics, we need better student-level studies that recruit entire classrooms of children so that we might observe how such peer dynamics unfold. Moreover, such studies would be most informative if the assignment of pre-k *and* the subsequent composition of the peer environment were exogenously determined. Because it seems likely that instructional dynamics could translate to changes in the peer composition of a classroom, any study in this area would benefit from careful measurement of the instructional environment. And although the implications for theory and practice in this area would be profound, we know of no modern studies in the U.S. that have achieved this type of design. (However, see: Araujo et al., 2016.)

Narrowing the Field of Inquiry for Dynamic Complementarity

We found little evidence to support dynamic complementarity, but we acknowledge that our analyses might not have included key school variables that could test for this. Nationwide, an historical challenge in ECE has been the lack of coherence between early childhood programs and public school systems (Stipek et al., 2017). A key goal of many education planners today is to "align" preschool and elementary school curricula. As we suggested in Chapter III, dynamic complementarity might occur in specific instances in which learning trajectories are essential to promote further learning in a given domain (see Baroody et al., 2022). Studying broader measures, like school quality, would hardly capture whether children are exposed to specific curricular sequencing in a way that might produce complementary effects. Thus, future research must pay closer attention to measuring the differentiation of content between children of different achievement levels to uncover evidence of complementarity in a given achievement domain (Connor & Morrison, 2016). It should be noted that if we were to find clear evidence of dynamic complementarity in instances where content had been appropriately targeted, we might still wonder whether the policy implications would lead to desirable school reforms. Calls for differentiating instruction to better meet the needs of pre-k attenders could lead to more instances of student tracking in early grades. After all, the most efficient way to differentiate content between pre-k attenders and nonattenders may be to simply separate them into different classrooms within the same school. Although this approach may blunt the catch-up effect, we are skeptical that it would actually help meet the broader goals of educational equity that led us to pursue these questions in the first place.

Further, the failure to find strong synergies between NC Pre-K and school characteristics does not mean that such synergies do not exist or could not be created with different policies that we were not able to evaluate here. One opportunity for inquiry comes with the North Carolina state legislature's creation of the Birth Through Third Grade Interagency Council to facilitate coordination between the NC Department of Health and Human Services (where many early childhood programs are housed) and the NC Department of Public Instruction (where K-12 education is housed; www.b3council.nc.gov). Another initiative, Pathways to Grade-Level Reading, seeks to ensure that all children in North Carolina have the preschool and elementary school experiences needed to achieve grade-level reading in third grade (www.buildthefoundation.org/initiative/pathways-to-grade-level-reading/). Rigorous evaluation of these initiatives could test the hypothesis of dynamic complementarity and inform policy.

Outcomes Beyond Grade School Academic Achievement

Limitations in the available administrative data sets prevented us from examining impacts of pre-k on other important outcomes such as school attendance, disciplinary infractions, delinquency, social-emotional learning, health, and well-being. Some of these variables (e.g., attendance, disciplinary infractions) had been poorly measured in North Carolina records during the period covered in this study but have since been improved and could be examined in future studies. Administrative records of other variables (e.g., health, pregnancy) might be available for future inquiry. Still other variables (e.g., social-emotional learning, well-being) might require original data collection in future studies but are well worth investigation given the broad goals of pre-kindergarten and all of education.

Additionally, the positive impacts identified in this study need to be monetized so that they may provide more utility for policy makers. Monetized gains likely accrue across the lifespan. As we wait for the children in the current study to become adults, shadow prices (i.e., the estimated lifelong monetary value of an early behavioral marker) could be computed for intermediate outcomes (i.e., the monetary benefit of higher test scores in Grade 5) that could leverage a benefit-cost analysis of NC Pre-K. However, some work has called into question the wisdom of using test score gains to project impacts on adult outcomes, such as earnings, since correlations between student characteristics and adult outcomes can fluctuate depending on a range of factors (Watts, 2020).

Thus, we simply need more examination of rigorous studies that track students into adulthood. This need has never been greater given the puzzling findings from studies such as Gray-Lobe et al. (2021) that have uncovered clear evidence of impacts on adult outcomes for access to preschool, even when medium-term test score impacts were not observed (see also Deming, 2009). Indeed, the enduring impact of studies such as the Perry Preschool and Abecedarian evaluations have been largely drawn from the fact that researchers continued to follow the samples well into adulthood, allowing for the measurement of key outcomes that have proven to be essential to weighing lifelong benefits against program costs. We hope that other pre-k evaluations can continue to follow their samples into the late adolescent and early adulthood years, as we simply need more longitudinal studies with rigorous research designs to better understand the connections between short-run impacts and adult outcomes.

Considering the Counterfactual

As we described in Chapter II, the experiences of children assigned to business-as-usual control groups have changed tremendously since the studies of Perry Preschool and Abecedarian, when low-income and minoritized families were systematically denied opportunities for early education. Opportunities have increased along with motivation by families to find enriching educational experiences. This increase in opportunity has led to challenges for modern preschool research. For example, Weiland and colleagues (2020) found that almost all of the families in the “control” condition enrolled in an alternative type of formal childcare. The test being conducted in many studies is no longer one of early education versus absent in-school education but, rather, a test of one form of educational experience versus another form. This trend has implications for future evaluation study designs, the interpretation of findings, and the formulation of ECE policy.

Conclusion

Greater funding for the NC Pre-K program improves children's academic achievement through the end of elementary school, but impacts were not detected for special education placement and grade retention. The achievement impacts for this program are strongest for children born into disadvantage. Funding for the program acts as a buffer against prior adversity and a protective factor against some forms of future adversity. The program largely supplements the positive impacts of other investments in an additive way, and those other investments and positive environments continue to be beneficial for children regardless of pre-k exposure. Because programs and contexts vary greatly, it is not clear whether other pre-kindergarten programs would have the same effects. Policymakers who want to maximize children's learning would be wise to continue to support funding for high-quality statewide pre-kindergarten without diminishing support for other educational experiences. In total, our findings suggest that such funding will be most effective if targeted at children who are most in need of high quality early educational experiences.

Acknowledgments

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 the relevant data files and guidance. We would also like to thank the members of the Consortium on Early Childhood Intervention Impact, including Drew Bailey, Douglas Clements, Eric Dearing, Greg Duncan, Kelley Durkin, Dale Farran, William Gormley, Julie Sarama, Deborah Vandell, and Christina Weiland for their helpful comments on previous versions of this work. We also thank Gayane Baziyants, Caroline Botvin, and Hunter Kaufman for their assistance in preparing the monograph, as well as Susan Kruglinski for her excellent editing work. All errors are the fault of the authors. This research study was supported by grant funding from the National Institute of Child Health and Human Development (1R01HD095930-01A1).

Notes

1. Throughout the manuscript, we refer to educational programs that serve children at age 4 as “pre-k,” which is short for “pre-kindergarten.” However, we recognize that in much of the literature we cite, these same programs are sometimes referred to as “preschool.” When possible, we are careful to note if programs discussed also targeted children at younger ages.

2. North Carolina's public preschool program was originally called “More at Four,” before being relabeled as “NC Pre-K” in later years. In the current monograph, we refer to the program as NC Pre-K, though our work does consider years when the program would have been called More at Four.

3. In some cases, funding was allocated to small groups of counties, and then split amongst the group. Our calculations reflect those divisions.

4. https://eclkc.ohs.acf.hhs.gov/federal-monitoring/report/agency-service-profile.

5. Additionally, the monotonicity assumption implies that the instrument “moves” units in the same direction, meaning there are no “defiers” of the causal chain. In this case, a defier would be a family that became less likely to enroll in NC Pre-K as a result of funding increases, which is very unlikely. Furthermore, in a framework of heterogeneous treatment effects in IV—the subject of our study—there is an additional assumption that the instrument is as good as randomly assigned, meaning independent of potential outcomes and treatment assignments (Angrist & Pischke, 2009). This has been demonstrated for NC Pre-K funding in prior published papers (see Ladd et al., 2014).

6. The interaction for low birthweight shown in Table 6 did produce a statistically significant effect, but it indicated that NC Pre-K effects were *larger* for low birthweight children. This effect is in the opposite direction of what is implied by the subgroup results shown in the main text Table 3 when a separate regression was run for each subgroup. We believe this discrepancy is likely due to the difference in the way control variables are handled between the two approaches. When running the interaction model, we only varied the effect of NC Pre-K for “low” and “normal/high” birthweight children. Thus, control variables were assumed to have an equal effect for both groups. When we split the sample based on birthweight grouping, this approach implicitly assumes that the control variable effects also differ between the two groups, as each control variable (and each fixed effect) is modeled separately for both groups. Indeed, we checked our findings against a fully-interacted model that included an interaction term for each control and fixed effect, and it produced a nonsignificant effect in line with the subgroup difference shown in Table 3. In sum, we do not believe the interaction or subgroup results warrant a strong theoretical interpretation for the birthweight variable given the small magnitude of the apparent difference and the conflicting results from the two approaches.

7. Note that unlike Ananat et al. (2011), we measured job losses across the entire year—corresponding to the year within which the child took their 5th-grade test (i.e., from July to the following June)—in contrast to their 2011 study which found results were localized to the quarter immediately preceding the test administration.

ORCID

Tyler W. Watts http://orcid.org/0000-0002-2741-0873

Jade M. Jenkins http://orcid.org/0000-0002-2000-3087

Kenneth A. Dodge http://orcid.org/0000-0001-5932-215X

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

Administration for Children and Families. (2021). *Head Start Program Facts Fiscal Year 2019*.

Ananat, E. O., Gassman-Pines, A., & Gibson-Davis, C. M. (2011). The effects of local employment losses on children's educational achievement. In G. J. Duncan & R. J. Murnane (Eds.), *Whither opportunity* (pp. 299–314). Russel Sage Foundation.

Angrist, J. D., & Pischke, J. S. (2008). *Mostly harmless econometrics: An empiricists companion*. Princeton University Press.

Ansari, A., & Pianta, R. C. (2018). The role of elementary school quality in the persistence of preschool effects. *Children and Youth Services Review*, *86*, 120–127. https://doi.org/10.1016/j.childyouth.2018.01.025

Ansari, A., Purtell, K., & Gershoff, E. (2016). Classroom age composition and the school readiness of 3-and 4-year-olds in the Head Start program. *Psychological Science*, *27*(1), 53–63.

Araujo, M. C., Carneiro, P., Cruz-Aguayo, Y., & Schady, N. (2016). Teacher quality and learning outcomes in kindergarten*. *The Quarterly Journal of Economics*, *131*(3), 1415–1453. https://doi.org/10.1093/qje/qjw016

Ashe, M. K., Reed, S., Dickinson, D. K., Morse, A. B., & Wilson, S. J. (2009). Opening the world of learning: Features, effectiveness, and implementation strategies. *Early Childhood Services: An Interdisciplinary Journal of Effectiveness*, *3*(3), 179–191.

Auger, A., Farkas, G., Burchinal, M. R., Duncan, G. J., & Vandell, D. L. (2014). Preschool center care quality effects on academic achievement: An instrumental variables analysis. *Developmental Psychology*, *50*(12), 2559–2571.

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*, 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., Duncan, G. J., Watts, T., Clements, D. H., & Sarama, J. (2018). Risky business: Correlation and causation in longitudinal studies of skill development. *American Psychologist*, *73*(1), 81–94. https://doi.org/10.1037/amp0000146

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*, 100910. https://doi.org/10.1016/j.dr.2020.100910

Barnett, W. S., Epstein, D. J., Carolan, M. E., Fitzgerald, J., Ackerman, D. J., & Friedman, A. H. (2010). *The State of Preschool 2010*. Rutgers Graduate School of Education.

Barnett, W. S., Friedman-Krauss, A. H., Weisenfeld, G. G., Horowitz, M., Kasmin, R., & Squires, J. H. (2017). *The State of Preschool 2016*. National Institute for Early Education Research.

Barnett, W. S., Hustedt, J. T., Robin, K. B., & Schulman, K. (2003). *The state of preschool*. National Institute for Early Education Research.

Baroody, A. J., Clements, D. H., & Sarama, J. (2022). Lessons learned from 10 experiments that tested the efficacy and assumptions of hypothetical learning trajectories. *Education Sciences*, *12*(3), 195. https://doi.org/10.3390/educsci12030195

Bassok, D. (2012). Competition or collaboration?: Head Start enrollment during the rapid expansion of state pre-kindergarten. *Educational Policy*, *26*(1), 96–116. https://doi.org/10.1177/0895904811428973

Bassok, D., Fitzpatrick, M. D., & Loeb, S. (2014). Does state preschool crowd-out private provision? The impact of universal preschool on the childcare sector in Oklahoma and Georgia. *Journal of Urban Economics*, *83*, 18–33. https://doi.org/10.1016/j.jue.2014.07.001

Bassok, D., Gibbs, C. R., & Latham, S. (2018). Preschool and children's outcomes in elementary school: Have patterns changed nationwide between 1998 and 2010. *Child Development*, *90*, 1875–1897. https://doi.org/10.1111/cdev.13067

Bassok, D., Miller, L. C., & Galdo, E. (2016). The effects of universal state pre-kindergarten on the child care sector: The case of Florida's voluntary Pre-kindergarten program. *Economics of Education Review*, *53*, 87–98. https://doi.org/10.1016/j.econedurev.2016.05.004

Baulos, A., & Heckman, J. (2022). The importance of investing in early childhood development and the role of families. In A. Betts & K. Thai (Eds.), *Handbook of research on innovative approaches to early childhood development and school readiness* (pp. 38–54). IGI Global.

Baum, C. F., Schaffer, M. E., & Stillman, S. (2010). *ivreg2: Stata module for extended instrumental variables/2SLS, GMM and AC/HAC, LIML and k-class regression*. http://ideas.repec.org/c/boc/bocode/s425401.html

Bierman, K. L., Domitrovich, C. E., Nix, R. L., Gest, S. D., Welsh, J. A., Greenberg, M. T., Blair, C., Nelson, K. E., & Gill, S. (2008). Promoting academic and social-emotional school readiness: The Head Start REDI Program. *Child Development*, *79*(6), 1802–1817. https://doi.org/10.1111/j.1467-8624.2008.01227.x

Bitler, M. P., Hoynes, H. W., & Domina, T. (2014). *Experimental evidence on distributional effects of Head Start* (National Bureau of Economic Research Working Paper Series, No. 20434). https://doi.org/10.3386/w20434

Blad, E. (2021, April 28). *Biden pitches plan to expand universal pre-K, free school meal programs, teacher training*.

Bloom, H. S., & Michalopoulos, C. (2013). When is the story in the subgroups? *Prevention Science*, *14*(2), 179–188. https://doi.org/10.1007/s11121-010-0198-x

Bloom, H. S., & Weiland, C. (2015). *Quantifying variation in Head Start effects on young children's cognitive and socio-emotional skills using data from the National Head Start Impact Study*. Available at SSRN 2594430.

Brooks-Gunn, J., Klebanov, P. K., Liaw, F.-r., & Spiker, D. (1993). Enhancing the development of low-birthweight, premature infants: Changes in cognition and behavior over the first three years. *Child Development*, *64*(3), 736–753. https://doi.org/10.1111/j.1467-8624.1993.tb02940.x

Brooks-Gunn, J., McCarton, C. M., Casey, P. H., McCormick, M. C., Bauer, C. R., Bernbaum, J. C., Tyson, J., Swanson, M., Bennett, F. C., Scott, D. T., Tonascia, J., Meinert, C. L., Casey, P., Pearson, K., Rickert, V., Whitmarsh-Barrett, K., McCarton, C. M., Hogan, A., Kenny, C., & Shapiro, S. (1994). Early intervention in low-birth-weight premature infants: Results through age 5 years from the Infant Health and Development Program. *JAMA*, *272*(16), 1257–1262. https://doi.org/10.1001/jama.1994.03520160041040

Brown, E. R., Lanfredi, G., de Silva, A., Janta, B., & Devaux, A. (2022). *Building a better understanding of the impact of early childhood education and care on medium- and long-term educational and labour market outcomes in Europe*. Prepared as part of the European Platform for Investing in Children (EPIC) project. Publications of the European Union. https://doi.org/10.2767/559483

Brown, T. T., & Jernigan, T. L. (2012). Brain development during the preschool years. *Neuropsychology Review*, *22*(4), 313–333.
Bruno, E. P., & Iruka, I. U. (2022). Reexamining the Carolina abecedarian project using an anti-racist perspective: Implications for early care and education research. *Early Childhood Research Quarterly*, *58*, 165–176. https://doi.org/10.1016/j.ecresq.2021.09.001
Burchinal, M., Pianta, R., Ansari, A., Whittaker, J., & Vitiello, V. (2023). Kindergarten academic and social skills and exposure to peers with pre-kindergarten experience. *Early Childhood Research Quarterly*, *62*, 41–52.
Burchinal, M., Whitaker, A. A., & Jenkins, J. M. (2022). The promise and purpose of early care and education. *Child Development Perspectives*, *16*, 134–140. https://doi.org/10.1111/cdep.12463
Bush, G. (1990). A*ddress Before a Joint Session of the Congress on the State of the Union, January 31, 1990*. Weekly Compilation of Presidential Documents, National Archives and Record Administration.
Bustamante, A. S., Dearing, E., Zachrisson, H. D., & Vandell, D. L. (2022). Adult outcomes of sustained high-quality early child care and education: Do they vary by family income? *Child Development*, *93*, 502–523. https://doi.org/10.1111/cdev.13696
Callaway, B., Goodman-Bacon, A., & Sant'Anna, P. H. (2021). Difference-in-differences with a continuous treatment. *arXiv preprint arXiv:2107.02637*.
Campbell, F. A., Conti, G., Heckman, J. J., Moon, S. H., Pinto, R., Pungello, E., & Pan, Y. (2014). Early childhood investments substantially boost adult health. *Science*, *343*(6178), 1478–1485. https://doi.org/10.1126/science.1248429
Campbell, F. A., Pungello, E., Burchinal, M. R., Kainz, K., Pan, Y., Wasik, B. H., Barbarin, O. A., Sparling, J. J., & Ramey, C. T. (2012). Adult outcomes as a function of an early childhood educational program: An Abecedarian Project follow-up. *Developmental Psychology*, *48*(4), 1033.
Campbell, F. A., Pungello, E., Miller-Johnson, S., Burchinal, M. R., & Ramey, C. T. (2001). The development of cognitive and academic abilities: Growth curves from an early childhood educational experiment. *Developmental Psychology*, *37*(2), 231–242.
Cascio, E. U. (2021). Early childhood education in the United States: What, when, where, who, how, and why 1. In B. P. McCall (Ed.), *The Routledge handbook of the economics of education* (pp. 30–72). Routledge.
Cascio, E. U., & Schanzenbach, D. W. (2013). The impacts of expanding access to high-quality preschool education. *Brookings Papers on Economic Activity*, *2013*, 127–178.
Chaudry, A., & Datta, R. (2017). The current landscape for public pre-kindergarten programs. In K. A. D. D. Phillips & Pre-Kindergarten Task Force, (Eds.), *The current state of scientific knowledge on pre-kindergarten effects* (pp. 5–18). Brookings Institution and Duke University.
Chetty, R., Hendren, N., & Katz, L. F. (2016). The effects of exposure to better neighborhoods on children: New evidence from the Moving to Opportunity experiment. *American Economic Review*, *106*(4), 855–902.
Child Care & Early Education Research Connections. (2011). Early care and education collaboration: A key topic resource list. Office of Planning, Research, and Evaluation. https://www.researchconnections.org/childcare/resources/22863/pdf
Claessens, A., Engel, M., & Curran, F. C. (2013). Academic content, student learning, and the persistence of preschool effects. *American Educational Research Journal*, *51*(2), 403–434. https://doi.org/10.3102/0002831213513634
Clements, D. H., & Sarama, J. (2007). Effects of a preschool mathematics curriculum: Summative research on the Building Blocks project. *Journal for Research in Mathematics Education*, *38*(2), 136–163.
Clements, D. H., Sarama, J., Spitler, M. E., Lange, A. A., & Wolfe, C. B. (2011). Mathematics learned by young children in an intervention based on learning trajectories: A large-scale cluster randomized trial. *Journal for Research in Mathematics Education*, *42*(2), 127–166.
Clements, D. H., Sarama, J., Wolfe, C. B., & Spitler, M. E. (2013). Longitudinal evaluation of a scale-up model for teaching mathematics with trajectories and technologies persistence of effects in the third year. *American Educational Research Journal*, *50*(4), 812–850.

Cobb, C. T. (2009). North Carolina's More at Four prekindergarten program: A case study of funding versus quality and other issues in large-scale implementation. In R. Pianta, & C. Howes (Eds.), *The promise of pre-k* (pp. 123–144). Paul H. Brookes Publishing Co.

Connor, C. M., & Morrison, F. J. (2016). Individualizing student instruction in reading: Implications for policy and practice. *Policy Insights from the Behavioral and Brain Sciences*, *3*(1), 54–61. https://doi.org/10.1177/2372732215624931

Connors-Tadros, L. (2019). Why collaboration is important now and opportunities for future research. *Early Education and Development*, *30*(8), 1094–1099. https://doi.org/10.1080/10409289.2019.1662689

Correia, S. (2017). *Linear models with high-dimensional fixed effects: An Efficient and feasible estimator*. Working paper. http://scorreia.com/research/hdfe.pdf

Cunha, F., & Heckman, J. J. (2007). The technology of skill formation. *American Economic Review*, *97*(2), 31–47.

Cunha, F., Heckman, J. J., Lochner, L., & Masterov, D. V. (2006). Interpreting the evidence on life cycle skill formation. In E. Hanushek & F. Welch (Eds.), *Handbook of the economics of education* (Vol. 1, pp. 697–812). Elsevier.

Cunningham, S. (2021). *Causal inference: The mixtape*. Yale University Press.

Currie, J., & Almond, D. (2011). Human capital development before age five. In D. Card & O. Ashenfelter (Eds.), *Handbook of labor economics* (Vol. 4, pp. 1315–1486). Elsevier.

Currie, J., & Thomas, D. (2000). School quality and the longer-term effects of Head Start. *Journal of Human Resources*, *35*(4), 755–774.

Davis-Kean, P. E., Tighe, L. A., & Waters, N. E. (2021). The role of parent educational attainment in parenting and children's development. *Current Directions in Psychological Science*, *30*(2), 186–192. https://doi.org/10.1177/0963721421993116

Deming, D. (2009). Early childhood intervention and life-cycle skill development: Evidence from Head Start. *American Economic Journal: Applied Economics*, *1*(3), 111–134. https://doi.org/10.2307/25760174

Dodge, K. A., Bai, Y., Ladd, H. F., & Muschkin, C. G. (2017). 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

Duer, J. K., & Jenkins, J. (2022). Paying for preschool: Who blends funding in early childhood education. *Educational Policy*. https://doi.org/10.1177/08959048221103804

Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K. A., Huston, A. C., Klebanov, P., Pagani, L. S., Feinstein, L., Engel, M., Brooks-Gunn, J., Sexton, H., Duckworth, K., & Japel, C. (2007). School readiness and later achievement. *Developmental Psychology*, *43*(6), 1428–1446. https://doi.org/10.1037/0012-1649.43.6.1428

Duncan, G. J., Kalil, A., Mogstad, M., & Rege, M. (2022). Investing in early childhood development in preschool and at home (29985). National Bureau of Economic Research.

Duncan, G. J., & Magnuson, K. (2013). Investing in preschool programs. *The Journal of Economic Perspectives*, *27*(2), 109–132. https://doi.org/10.1257/jep.27.2.109

Durkin, K., Lipsey, M. W., Farran, D. C., & Wiesen, S. E. (2022). Effects of a statewide prekindergarten program on children's achievement and behavior through sixth grade. *Developmental Psychology*, *58*(3), 470–484. https://doi.org/10.1037/dev0001301

Early, D. M., Maxwell, K. L., Burchinal, M. R., Alva, S., Bender, R. H., Bryant, D., Cai, K., Clifford, R. M., Ebanks, C., Griffin, J. A., Henry, G. T., Howes, C., Iriondo-Perez, J., Jeon, H. J., Mashburn, A. J., Peisner-Feinberg, E., Pianta, R. C., Vandergrift, N., & Zill, N. (2007). Teachers' education, classroom quality, and young children's academic skills: Results from seven studies of preschool programs. *Child Development*, *78*(2), 558–580. https://doi.org/10.1111/j.1467-8624.2007.01014.x

Education Commission of the States. (2014). *Kindergarten entrance age 50-state analysis*.

Elango, S., García, J. L., Heckman, J. J., & Hojman, A. (2016). Early childhood education. In R. A. Moffitt (Ed.), *Economics of means-tested transfer programs in the United States* (Vol. 2, pp. 235–297). University of Chicago Press. https://doi.org/10.3386/w21766

Engel, M., Claessens, A., & Finch, M. A. (2012). Teaching students what they already know? The (mis)alignment between mathematics instructional content and student knowledge in kindergarten. *Educational Evaluation and Policy Analysis*, *35*(2), 157–178. https://doi.org/10.3102/0162373712461850

Feller, A., Grindal, T., Miratrix, L., & Page, L. C. (2016). Compared to what? Variation in the impacts of early childhood education by alternative care type. *The Annals of Applied Statistics*, *10*(3), 1245–1285.

Fisher, P. A., Barker, T. V., & Blaisdell, K. N. (2020). A glass half full and half empty: The state of the science in early childhood prevention and intervention research. *Annual Review of Developmental Psychology*, *2*(1), 269–294. https://doi.org/10.1146/annurev-devpsych-121318-084958

Fitzpatrick, M. D. (2008). Starting school at four: The effect of universal pre-kindergarten on children's academic achievement. *The BE Journal of Economic Analysis & Policy*, *8*(1), 46.

Foster, E. M., & McLanahan, S. (1996). An illustration of the use of instrumental variables: Do neighborhood conditions affect a young person's chance of finishing high school? *Psychological Methods*, *1*(3), 249–260.

Friedman-Krauss, A., Barnett, W. S., Garver, K. A., Hodges, K. S., Weisenfeld, G. G., & Gardiner, B. A. (2021). *The State of Preschool 2020: State preschool yearbook*. National Institute for Early Education Research, Rutgers University.

Garces, E., Thomas, D., & Currie, J. (2002). Longer-term effects of Head Start. *The American Economic Review*, *92*, 999–1012.

Garcia, E., & Weiss, E. (2017). *Education inequalities at the school starting gate: Gaps, trends, and strategies to address them*. Economic Policy Institute.

García, J. L., Heckman, J. J., Leaf, D. E., & Prados, M. J. (2020). Quantifying the life-cycle benefits of an influential early-childhood program. *Journal of Political Economy*, *128*(7), 2502–2541. https://doi.org/10.1086/705718

Gassman-Pines, A., Ananat, E. O., & Gibson-Davis, C. M. (2014). Effects of statewide job losses on adolescent suicide-related behaviors. *American Journal of Public Health*, *104*(10), 1964–1970. https://doi.org/10.2105/ajph.2014.302081

Gennetian, L. A., Magnuson, K. A., & Morris, P. A. (2008). From statistical associations to causation: What developmentalists can learn from instrumental variables techniques coupled with experimental data. *Developmental Psychology*, *44*(2), 381–394.

Gilliam, W. S., & Ripple, C. H. (2004). What can be learned from state-funded preschool initiatives? A data-based approach to the Head Start devolution debate. In E. F. Zigler & S. J. Styfco (Eds.), *The Head Start debates* (pp. 477–497). Brookes Publishing.

Golberstein, E., Gonzales, G., & Meara, E. (2019). How do economic downturns affect the mental health of children? Evidence from the National Health Interview Survey. *Health Economics*, *28*(8), 955–970. https://doi.org/10.1002/hec.3885

Gormley, W. T. (2008). The effects of Oklahoma's pre-k program on Hispanic children. *Social Science Quarterly*, *89*(4), 916–936. https://doi.org/10.1111/j.1540-6237.2008.00591.x

Gormley, W. T., & Gayer, T. (2005). Promoting school readiness in Oklahoma: An evaluation of Tulsa's Pre-K program. *Journal of Human Resources*, *XL*(3), 533–558. https://doi.org/10.3368/jhr.XL.3.533

Gormley, W. T., Gayer, T., Phillips, D., & Dawson, B. (2005). The effects of universal Pre-K on cognitive development. *Developmental Psychology*, *41*(6), 872–884. https://doi.org/10.1037/0012-1649.41.6.872

Gormley, W. T., Phillips, D., & Anderson, S. (2018). The effects of Tulsa's Pre-K program on middle school student performance. *Journal of Policy Analysis and Management*, *37*(1), 63–87. https://doi.org/10.1002/pam.22023

Gray-Lobe, G., Pathak, P. A., & Walters, C. R. (2021, July 27–30). *The long-term effects of universal preschool in Boston* [presented at "SI 2021 Children"] (National Bureau of Economic Research Working Paper Series, No. 28756). https://doi.org/10.3386/w28756

Guryan, J., Hurst, E., & Kearney, M. (2008). Parental education and parental time with children. *Journal of Economic Perspectives*, *22*(3), 23–46.

Hanushek, E. A., Peterson, P. E., Talpey, L. M., & Woessmann, L. (2019). The achievement gap fails to close. *Education Next*, *19*(3), 8–17.

Harding, J. F., McCoy, D. C., & McCormick, M. P. (2020). Understanding alignment in children's early learning experiences: Policies and practices from across the United States. *Early Childhood Research Quarterly*, *52*, 1–4. https://doi.org/10.1016/j.ecresq.2019.12.007

Harding, J. F., Morris, P. A., & Hughes, D. (2015). The relationship between maternal education and children's academic outcomes: A theoretical framework. *Journal of Marriage and Family*, *77*(1), 60–76. https://doi.org/10.1111/jomf.12156

Heckman, J. J., Moon, S. H., Pinto, R., Savelyev, P. A., & Yavitz, A. (2010). The rate of return to the HighScope Perry Preschool Program. *Journal of Public Economics*, *94*(1), 114–128.

Hill, C. J., Gormley, W. T., & Adelstein, S. (2015). Do the short-term effects of a high-quality preschool program persist? *Early Childhood Research Quarterly*, *32*(3), 60–79.

Huang, F. L., Invernizzi, M. A., & Drake, E. A. (2012). The differential effects of preschool: Evidence from Virginia. *Early Childhood Research Quarterly*, *27*(1), 33–45. https://doi.org/10.1016/j.ecresq.2011.03.006

Huttenlocher, J., Levine, S., & Vevea, J. (1998). Environmental input and cognitive growth: A study using time-period comparisons. *Child Development*, *69*(4), 1012–1029.

Jenkins, J. M., Farkas, G., Duncan, G. J., Burchinal, M. R., & Vandell, D. (2016). Head Start at ages 3 and 4 versus Head Start followed by state pre-k: Which is more effective? *Educational Evaluation and Policy Analysis*, *38*(1), 88–112.

Jenkins, J. M., & Henry, G. T. (2016). Dispersed vs. centralized policy governance: The case of state early care and education policy. *Journal of Public Administration Research and Theory*, *26*(4), 709–725.

Jenkins, J. M., Watts, T. W., Magnuson, K., Gershoff, E. T., Clements, D. H., Sarama, J., & Duncan, G. J. (2018). Do high-quality kindergarten and first-grade classrooms mitigate preschool fadeout? *Journal of Research on Educational Effectiveness*, *11*(3), 339–374.

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.

Kang, C. Y., Duncan, G. J., Clements, D. H., Sarama, J., & Bailey, D. H. (2019). The roles of transfer of learning and forgetting in the persistence and fadeout of early childhood mathematics interventions. *Journal of Educational Psychology*, *111*(4), 590.

Kauerz, K. (2008). Learning from others: State efforts to expand services and build systems of early care and education. In A. R. Tarlov & M. P. Debbink (Eds.), *Investing in early childhood development: Evidence to support movement for educational change* (pp. 85–112). Palgrave Macmillan.

Kline, P., & Walters, C. (2016). Evaluating public programs with close substitutes: The case of Head Start. *Quarterly Journal of Economics*, *131*(4), 1795–1848.

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

Ladd, H. F., Muschkin, C. G., & Dodge, K. A. (2014). From birth to school: Early childhood initiatives and third-grade outcomes in North Carolina. *Journal of Policy Analysis and Management*, *33*(1), 162–187.

Lee, M.-J. (2005). *Micro-econometrics for policy, program, and treatment effects*. Oxford University Press.
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*, 155–176. https://doi.org/10.1016/j.ecresq.2018.03.005
List, J. A., Momeni, F., & Zenou, Y. (2020). *The social side of early human capital formation: Using a field experiment to estimate the causal impact of neighborhoods*. National Bureau of Economic Research.
Love, J. M., Chazan-Cohen, R., Raikes, H., & Brooks-Gunn, J. (2013). What makes a difference: Early Head Start evaluation findings in a developmental context. *Monographs of the Society for Research in Child Development*, *78*(1), vii–viii, 1–173.
Lucey, C., & Parti, T. (2021, July 6). Biden's Pre-K proposal faces questions from Republicans over federal role. *The Wall Street Journal*.
Ludwig, J., & Miller, D. L. (2007). Does Head Start improve children's life chances? Evidence from a regression discontinuity design. *The Quarterly Journal of Economics*, *122*(1), 159–208. https://doi.org/10.1162/qjec.122.1.159
Magnuson, K. A., Kelchen, R., Duncan, G. J., Schindler, H. S., Shager, H., & Yoshikawa, H. (2016). Do the effects of early childhood education programs differ by gender? A meta-analysis. *Early Childhood Research Quarterly*, *36*, 521–536. https://doi.org/10.1016/j.ecresq.2015.12.021
Magnuson, K. A., Ruhm, C., & Waldfogel, J. (2007). The persistence of preschool effects: Do subsequent classroom experiences matter? *Early Childhood Research Quarterly*, *22*(1), 18–38. https://doi.org/10.1016/j.ecresq.2006.10.002
Magnuson, K. A., & Waldfogel, J. (2008). *Steady gains and stalled progress: Inequality and the Black-White test score gap*. Russell Sage Foundation.
Mashburn, A. J., Pianta, R., Hamre, B. K., Downer, J. T., Barbarin, O. A., Bryant, D., Burchinal, M., Early, D. M., & Howes, C. (2008). Measures of classroom quality in prekindergarten and children's development of academic, language, and social skills. *Child Development*, *79*(3), 732–749. https://doi.org/10.1111/j.1467-8624.2008.01154.x
Mattera, S. K., Jacob, R., MacDowell, C., & Morris, P. A. (2021). Long-term effects of enhanced early childhood math instruction: The impacts of making pre-k count and high 5s on third-grade outcomes. *MDRC*.
McCartney, K., Dearing, E., Taylor, B. A., & Bub, K. L. (2007). Quality child care supports the achievement of low-income children: Direct and indirect pathways through caregiving and the home environment. *Journal of Applied Developmental Psychology*, *28*(5), 411–426. https://doi.org/10.1016/j.appdev.2007.06.010
McCormick, M. C., Brooks-Gunn, J., Buka, S. L., Goldman, J., Yu, J., Salganik, M., Scott, D. T., Bennett, F. C., Kay, L. L., Bernbaum, J. C., Bauer, C. R., Martin, C., Woods, E. R., Martin, A., & Casey, P. H. (2006). Early intervention in low birth weight premature infants: Results at 18 years of age for the Infant Health and Development Program. *Pediatrics*, *117*(3), 771–780. https://doi.org/10.1542/peds.2005-1316
McCormick, M. P., Weiland, C., Hsueh, J., Maier, M., Hagos, R., Snow, C., Leacock, N., & Schick, L. (2020). Promoting content-enriched alignment across the early grades: A study of policies & practices in the Boston Public Schools. *Early Childhood Research Quarterly*, *52*, 57–73. https://doi.org/10.1016/j.ecresq.2019.06.012
Miller, E. B., Farkas, G., Vandell, D. L., & Duncan, G. J. (2014). Do the effects of Head Start vary by parental preacademic stimulation? *Child Development*, *85*(4), 1385–1400.
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.
Muschkin, C. G., Ladd, H. F., Dodge, K. A., & Bai, Y. (2020). Gender differences in the impact of North Carolina's early care and education initiatives on student outcomes in elementary school. *Educational Policy*, *34*(2), 377–407.

National Academies of Sciences, Engineering, and Medicine. (2018). *Transforming the financing of early care and education* (E. P. Backes & L. R. Allen, Eds.). National Academies Press. https://books.google.com/books?id=YXllDwAAQBAJ

Nesbitt, K. T., & Farran, D. C. (2021). Effects of prekindergarten curricula: Tools of the mind as a case study. *Monographs of the Society for Research in Child Development*, *86*(1), 7–119. https://doi.org/10.1111/mono.12425

NICHD Early Child Care Research Network. (2001). Nonmaternal care and family factors in early development: An overview of the NICHD Study of Early Child Care. *Journal of Applied Developmental Psychology*, *22*(5), 457–492. https://doi.org/10.1016/S0193-3973(01)00092-2

North Carolina Division of Child Development and Early Education. (2011). *NC Pre-Kindergarten (NC Pre-K) program requirements*. https://ncchildcare.ncdhhs.gov/Portals/0/documents/pdf/N/ncpre-kprogramrequirements_08_15_2011.pdf

OECD. (2022). "How do early childhood education systems differ around the world?" In *Education at a Glance 2022: OECD Indicators*. OECD Publishing. https://doi.org/10.1787/63190ffa-en

Pages, R., Lukes, D. J., Bailey, D. H., & Duncan, G. J. (2020). Elusive Longer-run impacts of Head Start: Replications within and across cohorts. *Educational Evaluation and Policy Analysis*, *42*(4), 471–492. https://doi.org/10.3102/0162373720948884

Pearman, F. A. (2020). The moderating effect of neighborhood poverty on preschool effectiveness: Evidence from the Tennessee voluntary prekindergarten experiment. *American Educational Research Journal*, *57*(3), 1323–1357.

Pearman, F. A., Springer, M. P., Lipsey, M., Lachowicz, M., Swain, W., & Farran, D. (2020). Teachers, schools, and Pre-K effect persistence: An examination of the sustaining environment hypothesis. *Journal of Research on Educational Effectiveness*, *13*(4), 547–573. https://doi.org/10.1080/19345747.2020.1749740

Peisner-Feinberg, E. S. (2003). *Child and program characteristics of the North Carolina more at four pre-kindergarten program: Year 1 (January - June, 2002)*. Chapel Hill, NC: FPG Child Development Institute. Retrieved from https://fpg.unc.edu/sites/fpg.unc.edu/files/resource-files/MAF_Yr1_full_report_0.pdf

Peisner-Feinberg, E. S., Garwood, J. D., & Mokrova, I. (2017). *Children's outcomes through second grade: Findings from year 4 of Georgia's Pre-K longitudinal study*. FPG Child Develoment Institute, The University of North Carolina at Chapel Hill.

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

Peisner-Feinberg, E. S., Mokrova, I. L., & Anderson, T. L. (2017). *Effects of participation in the North Carolina Pre-Kindergarten Program at the end of kindergarten: 2015−2016 statewide evaluation*. FPG Child Development Institute.

Peisner-Feinberg, E. S., & Schaaf, J. (2010). *Long-term effects of the North Carolina more at four pre-kindergarten program: Children's reading and math skills at third grade*. Chapel Hill, NC: University of North Carolina, Frank Porter Graham Child Development Institute.

Peisner-Feinberg, E. S., & Schaaf, J. M. (2009). *Evaluation of the North Carolina more at four pre-kindergarten program: A look across time at children's outcomes and classroom quality from pre-k through kindergarten (2003−2009)*. FPG Child Development Institute, UNC at Chapel Hill.

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*. Chapel Hill: The University of North Carolina, FPG Child Development Institute.

Peisner-Feinberg, E. S., Schaaf, J. M., LaForett, D. R., Hildebrandt, L. M., & Sideris, J. (2014). *Effects of Georgia's Pre-K Program on children's school readiness skills: Findings from the 2012–2013 evaluation study*. 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., Gormley, W. T., & Lowenstein, A. E. (2009). Inside the pre-kindergarten door: Classroom climate and instructional time allocation in Tulsa's pre-K programs. *Early Childhood Research Quarterly*, *24*(3), 213–228. https://doi.org/10.1016/j.ecresq.2009.05.002

Phillips, D. A., & Pre-Kindergarten Task Force. (Eds.). (2017). *The Current state of scientific knowledge on pre-kindergarten effects*. Brookings Institution and Duke University.

Pianta, R., Barnett, W. S., Burchinal, M. R., & Thornburg, K. R. (2009). The effects of preschool education: What we know, how public policy is or is not aligned with the evidence base, and what we need to know. *Psychological Science in the Public Interest*, *10*(2), 49–88. https://doi.org/10.1177/1529100610381908

Pianta, R., Howes, C., Burchinal, M. R., Bryant, D., Clifford, R., Early, D. M., & Barbarin, O. A. (2005). Features of pre-kindergarten programs, classrooms, and teachers: Do they predict observed classroom quality and child-teacher interactions? *Applied Developmental Science*, *9*(3), 144–159. https://doi.org/10.1207/s1532480xads0903_2

Puma, M., Bell, S., Cook, R., & Heid, C. (2010). *Head Start Impact Study* (Final Report). US Department of Health and Human Services, Administration for Children and Families.

Puma, M., Bell, S., Cook, R., Heid, C., Broene, P., Jenkins, F., Mashburn, A., & Downer, J. T. (2012). *Third grade follow-up to the Head Start Impact Study: Final report* (OPRE Report 2012-45). Office of Planning, Research and Evaluation, Administration for Children and Families.

Purves, D., & Lichtman, J. W. (1985). *Principles of neural development*. Sinauer Associates Incorporated.

Ramey, C. T., & Ramey, S. L. (1998). Early intervention and early experience. *American Psychologist*, *53*(2), 109.

Raudenbush, S. W., & Bloom, H. S. (2015). Learning about and from a distribution of program impacts using multisite trials. *American Journal of Evaluation*, *36*(4), 475–499.

Raver, C. C., Jones, S. M., Li-Grining, C., Zhai, F., Bub, K., & Pressler, E. (2011). CSRP's impact on low-income preschoolers' preacademic skills: Self-regulation as a mediating mechanism. *Child Development*, *82*(1), 362–378.

Raver, C. C., Jones, S. M., Li-Grining, C. P., Metzger, M., Champion, K. M., & Sardin, L. (2008). Improving preschool classroom processes: Preliminary findings from a randomized trial implemented in Head Start settings. *Early Childhood Research Quarterly*, *23*(1), 10–26. https://doi.org/10.1016/j.ecresq.2007.09.001

Ritchie, S. J., & Bates, T. C. (2013). Enduring links from childhood mathematics and reading achievement to adult socioeconomic status. *Psychological Science*, *24*(7), 1301–1308.

Sabol, T. J., Hong, S. L. S., Pianta, R., & Burchinal, M. R. (2013). Can rating pre-K programs predict children's learning? *Science*, *341*(6148), 845–846.

Sameroff, A. J., & Haith, M. M. (1996). *The five to seven year shift*. The University of Chicago Press.

Schenke, K., Nguyen, T., Watts, T. W., Sarama, J., & Clements, D. H. (2017). Differential effects of the classroom on African American and non-African American's mathematics achievement. *Journal of Educational Psychology*, *109*(6), 794.

Schindler, H. S., McCoy, D. C., Fisher, P. A., & Shonkoff, J. P. (2019). A historical look at theories of change in early childhood education research. *Early Childhood Research Quarterly*, *48*, 146–154. https://doi.org/10.1016/j.ecresq.2019.03.004

Schweinhart, L. J., Montie, J., Xiang, Z., Barnett, W. S. S., Belfield, C. R., & Nores, M. (2005). *Lifetime effects: The High/Scope Perry Preschool Study through age 40*. High/Scope Press.

Stipek, D., Clements, D. H., Coburn, C., Franke, M., & Farran, D. C. (2017). PK-3: What does it mean for instruction? *SRCD Social Policy Report*, *30*(2), 2–22.

Swain, W. A., Springer, M. G., & Hofer, K. G. (2015). Early grade teacher effectiveness and pre-k effect persistence: Evidence from Tennessee. *AERA Open*, *1*(4). https://doi.org/10.1177/2332858415612751

Taie, M. (2014). Skill acquisition theory and its important concepts in SLA. *Theory and Practice in Language Studies*, *4*(9), 1971–1976.

Texas Education Agency. (2017). *Prekindergarten outcomes for Texas public school students*. https://tea.texas.gov/sites/default/files/Prekindergarten_Longitudinal_Study.pdf

Vandell, D. L. (2004). Early child care: The known and the unknown. *Merrill-Palmer Quarterly*, *50*(3), 387–414.

Vinovskis, M. A. (1999). *The road to Charlottesville: The 1989 education summit*. National Education Goals Panel. https://govinfo.library.unt.edu/negp/reports/negp30.pdf

Walters, C. (2015). Inputs in the production of early childhood human capital: Evidence from Head Start. *American Economic Journal: Applied Economics*, *7*(4), 76–102. https://doi.org/10.3386/w20639

Watts, T. W. (2020). Academic achievement and economic attainment: Reexamining associations between test scores and long-run earnings. *AERA Open*, *6*(2), 2332858420928985.

Watts, T. W., Duncan, G. J., Clements, D. H., & Sarama, J. (2018). What is the long-run impact of learning mathematics during preschool? *Child Development*, *89*(2), 539–555.

Watts, T. W., Duncan, G. J., & Rivas, M. (2019). *A reanalysis of impacts of the Tennessee Voluntary Prekindergarten program*. Annenberg Institute at Brown University.

Watts, T. W., Duncan, G. J., Siegler, R. S., & Davis-Kean, P. E. (2014). What's past is prologue: Relations between early mathematics knowledge and high school achievement. *Educational Researcher*, *43*(7), 352–360.

Weiland, C., Unterman, R., & Shapiro, A. (2021). The kindergarten hotspot: Literacy skill convergence between boston prekindergarten enrollees and nonenrollees. *Child Development*, *92*(2), 600–608. https://doi.org/10.1111/cdev.13499

Weiland, C., Unterman, R., Shapiro, A., Staszak, S., Rochester, S., & Martin, E. (2020). The effects of enrolling in oversubscribed prekindergarten programs through third grade. *Child Development*, *91*(5), 1401–1422.

Weiland, C., & Yoshikawa, H. (2013). Impacts of a prekindergarten program on children's mathematics, language, literacy, executive function, and emotional skills. *Child Development*, *84*(6), 2112–2130. https://doi.org/10.1111/cdev.12099

Whitehurst, G. J. (2018). *Does state pre-K improve children's achievement?* (Vol. 59). Brookings Institution. https://www.brookings.edu/wp-content/uploads/2018/07/Report3.pdf

Winship, C., & Morgan, S. L. (1999). The estimation of causal effects from observational data. *Annual Review of Sociology*, *25*, 659–706.

Wollons, R. (2000). *Kindergartens and cultures: The global diffusion of an idea*. Yale University Press.

Xue, Y., Miller, E. B., Auger, A., Pan, Y., Burchinal, M., Tien, H.-C., Peisner-Feinberg, E., Zaslow, M., & Tarullo, L. (2016). IV. Testing for dosage-outcome associations in early care and education. *Monographs of the Society for Research in Child Development*, *81*(2), 64–74. https://doi.org/10.1111/mono.12239

Yoshikawa, H., Weiland, C., Brooks-Gunn, J., Burchinal, M. R., Espinosa, L. M., Gormley, W. T., & Zaslow, M. J. (2013). *Investing in our future: The evidence base on preschool education*. Foundation for Child Development, Society for Research in Child Development.

Zhai, F., Brooks-Gunn, J., & Waldfogel, J. (2014). Head Start's impact is contingent on alternative type of care in comparison group. *Developmental Psychology*, *50*(12), 2572.

Authors

Tyler W. Watts, Ph.D., is an Assistant Professor in the Department of Human Development at Teachers College, Columbia University. He studies the long-term effects of educational interventions and conducts longitudinal studies of children's cognitive and social-emotional skill development.

Jade M. Jenkins, Ph.D., is an Associate Professor at the University of California Irvine School of Education studying early childhood policy. Her work is multidisciplinary, focusing on issues amenable to educational and social policy intervention, using diverse methods to evaluate programs and understand the mechanisms that promote child and family well-being.

Kenneth A. Dodge, Ph.D., is the William McDougall Distinguished Professor of Public Policy and Psychology and Neuroscience at Duke University. He studies the development, prevention, and public policy of problem behaviors in children. He has created Family Connects, a universal approach to family primary care in early life.

Robert C. Carr, Ph.D., is a Research Scientist at the FPG Child Development Institute, University of North Carolina at Chapel Hill. His research considers how early childhood programs and policies influence children's development of academic and social-behavioral skills, focusing on state-funded pre-kindergarten and Head Start preschool programming.

Maria Sauval is a Ph.D. Candidate in Education at the University of California, Irvine. She studies the effectiveness of policies to reduce poverty and inequality, with a focus on early childhood policies and their effects on children and families.

Yu Bai, Ph.D., is a Statistician at Duke University. His research interests include methodology, child development, education policy, health, and healthcare.

Maya Escueta is a Postdoctoral Associate at the Center for Child and Family Policy at Duke University. She works on the NICHD-funded Baby's First Years project and researches the effects of poverty alleviations policies on family dynamics and the early childhood home environment.

Jennifer Duer, Ph.D., is an Assistant Research Professor at the National Institute for Early Education Research (NIEER). Her research focuses on federal and state policies designed to improve the health and development of young children and their families with an emphasis on integrated approaches for comprehensive birth-to-5 services.

Helen Ladd is the Susan B. King Professor Emerita of Public Policy Studies and Economics at Duke University. Her research focuses on topics in education policy including pre-school education, school accountability, market-based reforms, school finance, and teacher labor markets. She is a past president of the Association for Public Policy Analysis and Management and a member of the National Academy of Education.

Clara Muschkin is an Associate Research Professor Emeritus of Public Policy at Duke University. She is a Sociologist and Demographer with an interdisciplinary research focus. In her studies, she asks how education policies that shape school environments can influence student behavior and academic performance.

Ellen Peisner-Feinberg, Ph.D., is a Research Professor Emerita and Senior Research Scientist Emerita at The University of North Carolina at Chapel Hill. Her research focuses on the concurrent and longitudinal effects of early educational experiences, with an emphasis on improving opportunities and reducing disparities through policy and practice.

Elizabeth Ananat (Ph.D. MIT 2006, MPP Michigan 2001) is the Mallya Professor of Women and Economics at Barnard College, Columbia University. She studies the intergenerational dynamics of poverty and inequality in the United States, with a focus on causal identification and on social policies.

Appendix

APPENDIX A: ADDITIONAL TABLES AND FIGURES

TABLE A1
DESCRIPTIVE STATISTICS OF MATH AND READING TEST SCORES BY ACADEMIC YEAR

	Math			Reading		
	Mean (1)	*SD* (2)	*N* (3)	Mean (4)	*SD* (5)	*N* (6)
1998	161.61	9.51	239	158.93	7.87	239
1999	160.27	9.49	6,088	55.08	7.96	45,970
2000	159.40	10.00	61,681	155.13	8.63	61,391
2001	259.84	9.52	64,027	155.78	8.17	63,644
2002	260.70	9.61	63,988	156.20	7.83	63,473
2003	262.20	8.77	63,988	256.84	7.93	63,744
2004	262.68	8.76	63,374	256.98	7.84	63,202
2005	261.95	9.46	64,220	257.16	7.73	63,996
2006	353.81	9.12	61,618	257.15	7.64	61,338
2007	354.54	9.15	62,244	257.61	7.58	61,981
2008	356.07	8.30	64,416	349.86	9.30	64,203
2009	356.91	8.21	66,184	351.75	8.29	65,867
2010	357.39	8.33	69,381	351.95	8.08	68,907
2011	357.67	8.39	73,586	352.20	8.08	73,057
2012	357.65	8.46	75,984	352.13	8.14	75,347
2013	449.72	9.33	74,835	449.59	9.38	74,137
2014	449.88	9.69	75,444	449.60	9.59	75,021
2015	449.91	10.01	70,498	449.00	10.27	70,458
2016	450.50	10.04	76,106	449.48	9.98	76,084
2017	443.80	9.62	10,060	442.96	10.44	10,059
2018	440.60	8.25	238	440.25	8.73	237

Note. The North Carolina Department of Public Instruction changed scoring criteria across years such that year-to-year changes in mean scores should not be interpreted as reflecting cohort differences.
SD = standard deviation.

TABLE A2

Pairwise Correlations of All Study Moderators

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1. Head Start 4-year-old saturation (%)	1.00																
2. Early Head Start 0–3-year-old saturation (%)	0.31	1.00															
3. Smart Start funding ($000s)	0.07	0.03	1.00														
4. HS grantee in county	0.10	0.14	−0.03	1.00													
5. 2003 HS 4-year-old enrollment (%)	0.49	0.26	0.00	0.01	1.00												
6. School-average achievement composite	−0.11	0.00	−0.07	0.12	−0.04	1.00											
7. Local PPE (in 2019 $000s)	−0.07	0.00	0.05	0.10	−0.05	0.44	1.00										
8. State PPE (in 2019 $000s)	0.17	−0.02	0.03	−0.35	0.19	−0.23	−0.15	1.00									
9. Federal PPE (in 2019 $000s)	0.22	0.04	0.45	−0.17	0.13	−0.53	−0.28	0.45	1.00								
10. Student–teacher ratio in the school	−0.07	−0.04	−0.10	0.17	−0.09	−0.13	−0.25	−0.44	−0.17	1.00							
11. National Board Certified teachers (%)	0.03	0.09	0.46	0.08	0.03	0.22	0.25	−0.11	0.20	−0.09	1.00						
12. Teachers with <3 years experience (%)	−0.07	−0.04	−0.08	0.13	−0.07	−0.28	−0.04	−0.16	−0.08	0.25	−0.27	1.00					
13. Annual teacher turnover (%)	−0.07	−0.10	−0.33	−0.00	−0.13	−0.30	−0.06	0.02	−0.14	0.11	−0.50	0.51	1.00				
14. Job loss (% affected)	−0.09	−0.04	−0.14	0.14	−0.06	0.05	−0.00	−0.10	−0.20	0.11	−0.25	0.12	0.14	1.00			
15. Median family income ($)	−0.26	−0.08	−0.28	0.16	−0.25	0.46	0.47	−0.49	−0.64	0.14	−0.09	0.13	0.13	0.13	1.00		
16. SNAP recipients (%)	0.15	0.05	0.48	−0.03	0.08	−0.58	−0.33	0.15	0.74	0.05	0.33	−0.01	−0.19	−0.19	−0.60	1.00	
17. Medicaid recipients (%)	0.16	0.01	0.27	−0.07	0.15	−0.70	−0.52	0.27	0.72	0.06	0.01	0.02	0.02	−0.06	−0.70	0.85	1.00
Observations	1,800																

Note. HS = Head Start; PPE = per-pupil expenditure; *SD* = standard deviation; SNAP = Supplemental Nutrition Assistance Program.

TABLE A3

First Stage Results Predicting NC Pre-K Enrollment (Full Output)

		Child's Race/Ethnicity				Gender		Birthweight		Mother Education	
	Full Sample (1)	Black (2)	Hispanic (3)	White (4)	Other Race (5)	Female (6)	Male (7)	Low (8)	Normal or High (9)	Less Than High School (10)	High School or More (11)
NC Pre-K	0.153**	0.184**	0.148**	0.122**	0.167**	0.154**	0.151**	0.167**	0.151**	0.155**	0.150**
funding ($000s)	(0.004)	(0.011)	(0.015)	(0.009)	(0.006)	(0.004)	(0.004)	(0.005)	(0.004)	(0.010)	(0.005)
Smart Start	−0.001	−0.002	0.005	−0.001	−0.002	−0.001	−0.002	−0.001	−0.001	−0.000	−0.001
funding ($000s)	(0.002)	(0.004)	(0.012)	(0.003)	(0.005)	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)
Male	0.001^{+}	0.000	0.000	0.002**	−0.000			0.002	0.001^{+}	−0.000	0.001*
	(0.001)	(0.001)	(0.003)	(0.000)	(0.002)			(0.002)	(0.001)	(0.001)	(0.001)
Child is Black	0.037**					0.035**	0.038**	0.040**	0.036**	0.026**	0.037**
	(0.004)					(0.005)	(0.005)	(0.010)	(0.004)	(0.006)	(0.004)
Child is Native	0.038**					0.041**	0.034**	0.027**	0.039**	0.036**	0.034**
American	(0.008)					(0.008)	(0.009)	(0.009)	(0.008)	(0.007)	(0.009)
Child is Asian	0.032**					0.032**	0.033**	0.014	0.034**	0.029*	0.026**
	(0.009)					(0.010)	(0.009)	(0.020)	(0.009)	(0.013)	(0.009)
Child is Hispanic	0.062**					0.058**	0.065**	0.047**	0.063**	0.013^{+}	0.081**
	(0.005)					(0.006)	(0.005)	(0.013)	(0.005)	(0.007)	(0.006)
Child is mixed race	0.027**					0.027**	0.026**	0.034**	0.026**	0.013**	0.027**
	(0.003)					(0.003)	(0.003)	(0.009)	(0.003)	(0.004)	(0.004)
Extremely low birth	0.014**	0.008	−0.013	0.014*	0.008	0.012*	0.017*			−0.010	0.021**
weight	(0.005)	(0.006)	(0.022)	(0.007)	(0.018)	(0.006)	(0.007)			(0.011)	(0.006)
Very low birth weight	0.009**	0.002	−0.022	0.018**	0.007	0.009*	0.009*			0.003	0.011**
	(0.003)	(0.005)	(0.019)	(0.004)	(0.016)	(0.004)	(0.004)			(0.006)	(0.004)
Low birth weight	0.005**	0.001	−0.003	0.010**	0.005	0.005**	0.006**			−0.001	0.007**

(*Continued*)

TABLE A3. (*Continued*)

		Child's Race/Ethnicity				Gender		Birthweight		Mother Education	
	Full Sample (1)	Black (2)	Hispanic (3)	White (4)	Other Race (5)	Female (6)	Male (7)	Low (8)	Normal or High (9)	Less Than High School (10)	High School or More (11)
	(0.001)	(0.001)	(0.005)	(0.001)	(0.005)	(0.001)	(0.001)			(0.002)	(0.001)
High birth weight	−0.000	0.003^{+}	−0.004	−0.002**	−0.004	0.000	−0.001			0.000	−0.000
	(0.001)	(0.001)	(0.004)	(0.001)	(0.003)	(0.001)	(0.001)			(0.002)	(0.001)
Mother's education	0.004^{+}	−0.005*	−0.018**	0.010**	0.001	0.004^{+}	0.003^{+}	0.002	0.004^{+}		
<12 years (1 = Yes)	(0.002)	(0.002)	(0.005)	(0.002)	(0.011)	(0.002)	(0.002)	(0.003)	(0.002)		
Mother marital status	−0.023**	−0.015**	−0.000	−0.028**	−0.019**	−0.023**	−0.023**	−0.018**	−0.023**	0.006**	−0.030**
	(0.002)	(0.003)	(0.005)	(0.001)	(0.005)	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)
Mother immigration	0.027**	0.013^{+}	0.062**	0.012**	0.014**	0.028**	0.026**	0.021*	0.028**	0.020**	0.030**
	(0.005)	(0.007)	(0.011)	(0.004)	(0.005)	(0.005)	(0.005)	(0.010)	(0.005)	(0.007)	(0.004)
Mother age in year	−0.002**	−0.001**	−0.001	−0.003**	−0.002**	−0.002**	−0.002**	−0.002**	−0.002**	−0.000**	−0.003**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
No father information	−0.019**	−0.005**	−0.023**	−0.003*	−0.008**	−0.018**	−0.019**	−0.015**	−0.019**	−0.011**	−0.018**
	(0.001)	(0.001)	(0.007)	(0.002)	(0.003)	(0.002)	(0.002)	(0.003)	(0.001)	(0.002)	(0.002)
First born	−0.018**	−0.008**	0.002	−0.022**	−0.017**	−0.018**	−0.017**	−0.016**	−0.018**	0.000	−0.019**
	(0.001)	(0.002)	(0.004)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)	(0.002)	(0.001)
Mother Black	0.019**	0.017**	0.041**	0.026**	0.029**	0.020**	0.017**	0.017^{+}	0.019**	0.011*	0.021**
	(0.004)	(0.005)	(0.013)	(0.009)	(0.006)	(0.004)	(0.006)	(0.009)	(0.004)	(0.006)	(0.004)
Mother Native	0.017*	0.012	0.054*	0.017	0.008	0.012	0.022**	0.013	0.017*	0.004	0.023**
American	(0.007)	(0.012)	(0.027)	(0.011)	(0.006)	(0.009)	(0.006)	(0.012)	(0.007)	(0.009)	(0.006)
Mother Asian	−0.036**	−0.039**	-0.044^{+}	0.003	−0.023**	−0.037**	−0.035**	-0.034^{+}	−0.036**	−0.040**	−0.034**
	(0.005)	(0.014)	(0.024)	(0.008)	(0.007)	(0.005)	(0.006)	(0.019)	(0.005)	(0.008)	(0.005)

(*Continued*)

TABLE A3. (*Continued*)

		Child's Race/Ethnicity				Gender		Birthweight		Mother Education	
	Full Sample (1)	Black (2)	Hispanic (3)	White (4)	Other Race (5)	Female (6)	Male (7)	Low (8)	Normal or High (9)	Less Than High School (10)	High School or More (11)
Mother Hispanic	0.036**	0.007	0.024**	0.019**	0.011	0.039**	0.033**	0.028*	0.037**	0.027**	0.041**
	(0.005)	(0.013)	(0.005)	(0.006)	(0.009)	(0.006)	(0.005)	(0.012)	(0.005)	(0.007)	(0.006)
Mother other race	−0.022*	0.041^{+}	−0.063**	0.032	−0.001	−0.021*	−0.024	−0.029	-0.022^{+}	−0.030*	−0.012
	(0.011)	(0.025)	(0.011)	(0.023)	(0.012)	(0.010)	(0.015)	(0.026)	(0.011)	(0.015)	(0.010)
% births from Black mothers in county/year	−0.000	0.001	0.001	0.000	0.000	−0.000	−0.000	−0.001	−0.000	0.002^{+}	-0.001^{+}
	(0.000)	(0.001)	(0.002)	(0.001)	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)	(0.001)	(0.000)
% births from Hispanic mothers in county/year	−0.000**	−0.000	−0.000	−0.000**	−0.000	−0.000**	−0.000**	−0.000*	−0.000**	−0.000**	−0.000**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
% births from mothers <12 years education county/year	−0.000	−0.000	−0.001	−0.003**	-0.002^{+}	−0.001	−0.000	−0.000	−0.001	−0.001	−0.000
	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)	(0.001)	(0.000)
Number of births in county and year (log)	−0.039**	−0.024	−0.092	−0.041*	−0.004	−0.033**	−0.045**	−0.039	−0.039**	-0.044^{+}	−0.032*
	(0.011)	(0.027)	(0.059)	(0.018)	(0.033)	(0.011)	(0.012)	(0.024)	(0.011)	(0.023)	(0.013)
County population (log)	0.048^{+}	0.128	0.500**	−0.009	0.106	0.057^{+}	0.038	0.041	0.049^{+}	0.221**	0.007
	(0.027)	(0.079)	(0.150)	(0.044)	(0.086)	(0.031)	(0.025)	(0.048)	(0.027)	(0.071)	(0.035)
Median family income in the year of birth	−0.000	−0.000	0.000	0.000	−0.000*	−0.000	−0.000	0.000	−0.000	0.000	−0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

(*Continued*)

TABLE A3. (*Continued*)

		Child's Race/Ethnicity				Gender		Birthweight		Mother Education	
	Full Sample (1)	Black (2)	Hispanic (3)	White (4)	Other Race (5)	Female (6)	Male (7)	Low (8)	Normal or High (9)	Less Than High School (10)	High School or More (11)
% SNAP recipients in the year of birth	−0.002	0.004	0.000	−0.007**	−0.003	−0.002	−0.002	−0.000	−0.002	−0.000	−0.002
	(0.001)	(0.002)	(0.003)	(0.002)	(0.002)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.001)
% Medicaid population in the year of birth	0.001^{+}	−0.003	0.000	0.006**	0.007**	0.002*	0.001	−0.002	0.002*	−0.001	0.002
	(0.001)	(0.002)	(0.005)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)
Observations	1,200,135	346,630	81,686	703,248	68,570	597,325	602,810	96,856	1,103,081	280,345	919,790
Enrollment mean	0.085	0.113	0.246	0.048	0.127	0.085	0.086	0.103	0.084	0.125	0.073

Note. NC Pre-K = North Carolina Pre-K; SNAP = Supplemental Nutrition Assistance Program.
$^{+}p < .10$.
$^{*}p < .05$.
$^{**}p < .01$.
$^{***}p < .001$.

TABLE A4

IV Results for the Effect of NC Pre-K Enrollment on 5th Grade Reading and Math Achievement

		Child's Race/Ethnicity				Gender		Birthweight		Mother Education		Economic Disadvantage	
	Full Sample (1)	Black (2)	His- panic (3)	White (4)	Other (5)	Female (6)	Male (7)	Low (8)	Normal or High (9)	Less Than High School (10)	High School or More (11)	No (12)	Yes (13)
Reading													
NC Pre-K enrollment	0.219**	0.167*	0.272**	0.206**	0.085	0.204**	0.236**	0.153	0.226**	0.326**	0.164*	0.039	0.193**
	(0.063)	(0.076)	(0.085)	(0.065)	(0.105)	(0.067)	(0.066)	(0.099)	(0.061)	(0.076)	(0.064)	(0.094)	(0.067)
Observations	1,200,135	346,630	81,686	703,248	68,570	597,325	602,810	96,856	1,103,081	280,345	919,790	597,914	584,481
F statistic of instrument	1,470.089	264.026	92.675	184.153	702.227	1,243.040	1,486.708	980.066	1,405.453	226.806	764.630	245.349	509.583
Math													
NC Pre-K enrollment	0.158*	0.171^{+}	0.285*	0.069	0.057	0.185*	0.131	0.132	0.159*	0.233**	0.108	−0.057	0.148^{+}
	(0.078)	(0.090)	(0.109)	(0.086)	(0.121)	(0.079)	(0.083)	(0.104)	(0.078)	(0.080)	(0.084)	(0.100)	(0.084)
Observations	1,205,965	348,703	82,455	705,842	68,964	598,569	607,396	97,384	1,108,381	282,855	923,110	599,293	588,950
F statistic of instrument	1,462.263	264.467	92.908	182.819	704.669	1,238.136	1,473.594	974.218	1,403.737	227.404	766.764	244.395	511.281

Note. IV = instrumental variable; NC Pre-K = North Carolina Pre-K.
$^{+}p < .10$.
$^{*}p < .05$.
$^{**}p < .01$.
$^{***}p < .001$.

TABLE A5

IV Falsification Test With Subgroups by Economic Disadvantage (ED)

	First Stage NC Pre-K Enrollment		Second Stage Academic Achievement	
	IV—Not ED (1)	IV—ED (2)	IV—Not ED (3)	IV—ED (4)
NC Pre-K funding ($000s)	0.122***	0.161***		
	(0.008)	(0.007)		
NC Pre-K enrollment	−0.012	0.180*		
	(0.098)	(0.078)		
Observations	599,714	590,023	599,714	590,023
Enrollment mean	0.038	0.136		
F statistic of instrument	244.924	511.611		

Note. IV = instrumental variable; NC Pre-K = North Carolina Pre-K.
$^{+}p < .10$.
$^{*}p < .05$.
$^{**}p < .01$.
$^{***}p < .001$.

TABLE A6

Effects of NC Pre-K Funding on Reading Achievement: Moderation by the Early Childhood Environment

	Baseline With Chapter Subsample	Moderator: 4-Year-Old HS Enrollment		Moderator: EHS Enrollment (Ages 0–3)		Moderator: Smart Start			Mod: HS Indicator	Mod: 2003 HS Enroll.
		Main Effect	Interaction	Main Effect	Interaction	Main Effect	Interaction	All Mods.	Int. Only	Int. Only
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
NC Pre-K funding ($000s)	0.033***	0.034***	0.041***	0.034***	0.035***	0.033***	0.025*	0.032**	0.035***	0.033**
	(0.009)	(0.009)	(0.010)	(0.009)	(0.009)	(0.009)	(0.011)	(0.011)	(0.008)	(0.010)
HS saturation (in 10pp)		−0.005	−0.001					−0.003		
		(0.006)	(0.006)					(0.006)		
HS saturation × NC Pre-K funding			-0.012^{+}					−0.008		
			(0.007)					(0.008)		
Max EHS enrollment (in 10pp)				−0.064*	−0.049			−0.050		
				(0.029)	(0.031)			(0.031)		
Max EHS enrollment × NC Pre-K funding					−0.021			−0.012		
					(0.021)			(0.023)		
Smart Start funding ($000s)						0.006	0.007	0.006		
						(0.005)	(0.005)	(0.005)		
Smart Start funding × NC Pre-K funding							0.012	0.010		
							(0.009)	(0.009)		
HS grantee in county × NC Pre-K funding									−0.002	
									(0.010)	

(*Continued*)

TABLE A6. (*Continued*)

	Baseline With Chapter Subsample (1)	Moderator: 4-Year-Old HS Enrollment		Moderator: EHS Enrollment (Ages 0–3)		Moderator: Smart Start		All Mods. (8)	Mod: HS Indicator Int. Only (9)	Mod: 2003 HS Enroll. Int. Only (10)
		Main Effect (2)	Interaction (3)	Main Effect (4)	Interaction (5)	Main Effect (6)	Interaction (7)			
2003 HS 4-year-old enrollment × NC Pre-K funding										0.000 (0.001)
Observations	1,200,135	1,200,135	1,200,135	1,200,135	1,200,135	1,200,135	1,200,135	1,200,135	1,200,135	1,200,135
p value (interaction = 0)								.321		

Note. Standard errors clustered at the county level. All models include Smart Start funding (in $000s), child and family covariates, county covariates, and county, program year, and quarter of birth fixed effects. See Table 3 for details. No main effect models or main effects are shown for time-invariant measures of 2003 HS enrollment and the HS grantee in county indicator; these variables drop out in county fixed effects regression and can only be recovered through their interaction with a time-varying county feature (i.e., NC Pre-K funding).

EHS = Early Head Start; HS = Head Start; NC Pre-K = North Carolina Pre-K.

$^{+}p < .10$.
$^{*}p < .05$.
$^{**}p < .01$.
$^{***}p < .001$.

TABLE A7

Effects of NC Pre-K Funding on Math Achievement: Moderation by the Early Childhood Environment

	Baseline With Chapter Sub-sample	Moderator: 4-Year-Old HS Enrollment		Moderator: EHS Enrollment (Ages 0–3)		Moderator: Smart Start			HS as a "springboard" for NC Pre-K: Moderator: HS Indicator Variable	HS as a "springboard" for NC Pre-K: Moderator: 2003 HS Enrollment
		Main Effect	Interaction	Main Effect	Interaction	Main Effect	Interaction	All Mods.	Interaction Only	Interaction Only
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
NC Pre-K funding ($000s)	0.024*	0.024*	0.031*	0.025*	0.027*	0.024*	0.017	0.024	0.026*	0.026*
	(0.012)	(0.012)	(0.013)	(0.011)	(0.012)	(0.012)	(0.014)	(0.015)	(0.011)	(0.013)
HS saturation (in 10pp)		−0.009	−0.005					−0.007		
		(0.008)	(0.008)					(0.008)		
HS saturation (in 10pp) × NC Pre-K funding			−0.012					−0.007		
			(0.009)					(0.010)		
Max EHS enrollment (in 10pp)				−0.086**	−0.063*			−0.063*		
				(0.029)	(0.029)			(0.029)		
Max EHS enrollment (in 10pp) × NC Pre-K funding					−0.032			−0.026		
					(0.023)			(0.025)		
Smart Start funding ($000s)						−0.003	−0.002	−0.003		
						(0.008)	(0.008)	(0.008)		
Smart Start funding ($000s) × NC Pre-K funding							0.009	0.008		
							(0.011)	(0.011)		

(Continued)

TABLE A7. (*Continued*)

	Baseline With Chapter Sub-sample (1)	Moderator: 4-Year-Old HS Enrollment — Main Effect (2)	Moderator: 4-Year-Old HS Enrollment — Inter-action (3)	Moderator: EHS Enrollment (Ages 0–3) — Main Effect (4)	Moderator: EHS Enrollment (Ages 0–3) — Inter-action (5)	Moderator: Smart Start — Main Effect (6)	Moderator: Smart Start — Inter-action (7)	All Mods. (8)	HS as a “springboard” for NC Pre-K — Modera-tor: HS Indicator Variable Interac-tion Only (9)	HS as a “springboard” for NC Pre-K — Moderator: 2003 HS Enrollment Interaction Only (10)
HS grantee in county × NC Pre-K funding									−0.003 (0.015)	
2003 HS 4-year-old enrollment × NC Pre-K funding										−0.001 (0.001)
Observations	1,205,965	1,205,965	1,205,965	1,205,965	1,205,965	1,205,965	1,205,965	1,205,965	1,205,965	1,205,965
p Value (interactions = 0)								.435		

Note. Standard errors clustered at the county level. All models include Smart Start funding (in $000s), child and family covariates, county covariates, and county, program year, and quarter of birth fixed effects. See Table 3. for details. No main effect models or main effects are shown for time-invariant measures of 2003 HS enrollment and the HS grantee in county indicator; these variables drop out in county fixed effects regression and can only be recovered through their interaction with a time-varying county feature (i.e., NC Pre-K funding).
EHS = Early Head Start; HS = Head Start; NC Pre-K = North Carolina Pre-K.
$^{+}p < .10$.
$^{*}p < .05$.
$^{**}p < .01$.
$^{***}p < .001$.

TABLE A8

Effects of NC Pre-K Funding on Reading Achievement: Moderation by Elementary School Characteristics

	Baseline With Chapter Sub-Sample	Moderator: (Lagged) Academic Composite		Moderator: Local PPE		Moderator: State PPE		Moderator: Federal PPE		
		Main Effect	Interaction	Main Effect	Interaction	Main Effect	Interaction	Main Effect	Interaction	All Mods.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
NC Pre-K funding ($000s)	0.033***	0.028***	0.024***	0.033***	0.035***	0.033***	0.033**	0.032***	0.018+	0.013
	(0.009)	(0.008)	(0.007)	(0.009)	(0.010)	(0.009)	(0.010)	(0.009)	(0.009)	(0.009)
School achievement composite		0.155***	0.157***							0.155***
(lagged;std.)		(0.004)	−0.004)							(0.004)
Sch. Ach. × NC Pre-K funding			−0.006*							−0.002
			(0.003)							(0.003)
Local PPE (in 2019 $000s)				0.050***	0.048***					0.017**
				(0.007)	(0.007)					(0.005)
Local PPE × NC Pre-K funding					0.003					0.007
					(0.008)					(0.007)
State PPE (in 2019 $000s)						−0.021**	−0.021**			0.037***
						(0.007)	−0.007			(0.008)
State PPE × NC Pre-K funding							0.000			−0.014*
							(0.005)			(0.006)
Federal PPE (in 2019 $000s)								−0.103***	−0.120***	−0.046*
								(0.018)	(0.020)	(0.019)
Federal PPE × NC Pre-K funding									0.030*	0.052***
									(0.013)	(0.014)
Observations	1,141,934	1,141,934	1,141,934	1,141,934	1,141,934	1,141,934	1,141,934	1,141,934	1,141,934	1,141,934
p Value (interactions = 0)										.003

Note. Standard errors are given in parentheses. Standard errors clustered at the county level. All models include Smart Start funding (in $000s), child and family covariates, county covariates, and county, program year, and quarter fixed effects. See Table 3 for details.
NC Pre-K = North Carolina Pre-K; PPE = per-pupil expenditure.
$^{+}p < .10$.
$^{*}p < .05$.
$^{**}p < .01$.
$^{***}p < .001$.

TABLE A9

EFFECTS OF NC PRE-K FUNDING ON MATH ACHIEVEMENT: MODERATION BY ELEMENTARY SCHOOL CHARACTERISTICS

	Baseline With Chapter Subsample	Moderator: (Lagged) Academic Composite		Moderator: Local PPE		Moderator: State PPE		Moderator: Federal PPE		
		Main Effect	Interaction	Main Effect	Interaction	Main Effect	Interaction	Main Effect	Interaction	All Mods.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
NC Pre-K funding ($000s)	0.023^{+}	0.017^{+}	0.010	0.023^{+}	0.018	0.023*	0.024^{+}	0.021^{+}	0.008	−0.004
	(0.012)	(0.009)	(0.009)	(0.012)	(0.012)	(0.012)	(0.013)	(0.011)	(0.012)	(0.010)
School-average achievement composite (lagged;std.)		0.187***	0.192***							0.190***
		(0.003)	(0.003)							(0.003)
School-average achievement composite × NC Pre-K funding			−0.011**							−0.006
			(0.004)							(0.004)
Local PPE (in 2019 $000s)				0.060***	0.063***					0.026**
				(0.007)	(0.009)					(0.008)
Local PPE × NC Pre-K funding					−0.009					−0.006
					(0.010)					(0.008)
State PPE (in 2019 $000s)						−0.033***	−0.031**			0.027**
						(0.009)	(0.010)			(0.008)

(*Continued*)

TABLE A9. (*Continued*)

	Baseline With Chapter Subsample	Moderator: (Lagged) Academic Composite		Moderator: Local PPE		Moderator: State PPE		Moderator: Federal PPE		
		Main Effect	Interaction	Main Effect	Interaction	Main Effect	Interaction	Main Effect	Interaction	All Mods.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
State PPE × NC Pre-K funding							−0.002			−0.018*
							(0.007)			(0.008)
Federal PPE (in 2019 $000s)								−0.103***	−0.119***	−0.010
								(0.020)	(0.023)	(0.021)
Federal PPE × NC Pre-K funding									0.030*	0.044*
									(0.014)	(0.018)
Observations	1,147,567	1,147,567	1,147,567	1,147,567	1,147,567	1,147,567	1,147,567	1,147,567	1,147,567	1,147,567
p Value (interactions = 0)										.081

Note. Standard errors are given in parentheses. Standard errors clustered at the county level. All models include Smart Start funding (in $000s), child and family covariates, county covariates, and county, program year, and quarter fixed effects. See Table 3 for details.
NC Pre-K = North Carolina Pre-K; PPE = per-pupil expenditure.
+$p < .10$.
*$p < .05$.
**$p < .01$.
***$p < .001$.

TABLE A10

Effects of NC Pre-K Funding on Reading Achievement: Moderation by Teacher Characteristics

	Baseline With Chapter Subsample	Moderator: Student–Teacher Ratio		Moderator: National Board Certification		Moderator: Teacher Experience		Moderator: Teacher Turnover		All Mods.
		Main Effect	Interaction	Main Effect	Interaction	Main Effect	Interaction	Main Effect	Interaction	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
NC Pre-K funding ($000s)	0.032***	0.032***	0.032**	0.034***	0.034***	0.030***	0.029***	0.031**	0.030**	0.030**
	(0.009)	(0.009)	(0.010)	(0.009)	(0.009)	(0.008)	(0.009)	(0.009)	(0.009)	(0.009)
Student–teacher ratio in the school		0.008***	0.007***							0.006***
		(0.001)	(0.001)							(0.001)
Student–teacher ratio in the school × NC Pre-K funding			0.004							0.003
			(0.003)							(0.002)
National Board Certified teachers (in 10pp units)				0.040***	0.039***					0.020**
				(0.006)	(0.008)					(0.006)
National Board Certified teachers × NC Pre-K funding					0.002					0.005
					(0.006)					(0.006)
Teachers with <3 years experience (in 10pp units)						−0.043***	−0.045***			−0.038***
						(0.003)	(0.004)			(0.004)

(*Continued*)

TABLE A10. (*Continued*)

	Baseline With Chapter Subsample	Moderator: Student–Teacher Ratio		Moderator: National Board Certification		Moderator: Teacher Experience		Moderator: Teacher Turnover		All Mods.
		Main Effect	Interaction	Main Effect	Interaction	Main Effect	Interaction	Main Effect	Interaction	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Teachers with <3 years experience × NC Pre-K funding							0.004			0.008**
							(0.003)			(0.003)
Annual teacher turnover (in 10pp units)								−0.030***	−0.029***	−0.011***
								(0.003)	(0.003)	(0.002)
Annual teacher turnover × NC Pre-K funding									−0.003	−0.000
									(0.003)	(0.003)
Observations	1,180,217	1,180,217	1,180,217	1,180,217	1,180,217	1,180,217	1,180,217	1,180,217	1,180,217	1,180,217
p Value (interactions = 0)										.074

Note. Standard errors are given in parentheses. Standard errors clustered at the county level. All models include Smart Start funding (in $000s), child and family covariates, county covariates, school-district covariates, and county, program year, and quarter fixed effects. See Table 3 for details. School district covariates: federal, state, and local per pupil expenditures in the school district.

NC Pre-K = North Carolina Pre-K.

$^{+}p < .10$.

$^{*}p < .01$.

$^{**}p < .05$.

$^{***}p < .001$.

TABLE A11

Effects of NC Pre-K Funding on Math Achievement: Moderation by Teacher Characteristics

	Baseline With Chapter Subsample	Moderator: Student–Teacher Ratio		Moderator: National Board Certification		Moderator: Teacher Experience		Moderator: Teacher Turnover		
		Main Effect	Interaction	Main Effect	Interaction	Main Effect	Interaction	Main Effect	Interaction	All Mods.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
NC Pre-K funding ($000s)	0.023*	0.023*	0.023^{+}	0.025*	0.025*	0.021^{+}	0.020^{+}	0.022^{+}	0.023^{+}	0.022^{+}
	(0.011)	(0.012)	(0.012)	(0.011)	(0.012)	(0.011)	(0.011)	(0.012)	(0.012)	(0.012)
Student–teacher ratio in the school		0.007***	0.007***							0.006***
		(0.001)	(0.001)							(0.001)
Student–teacher ratio in the school × NC Pre-K funding			0.001							−0.000
			(0.003)							(0.002)
National Board Certified teachers (in 10pp units)				0.045***	0.046***					0.025***
				(0.005)	(0.007)					(0.006)
National Board Certified teachers × NC Pre-K funding					−0.002					0.006
					(0.007)					(0.006)
Teachers with <3 years experience (in 10pp units)						−0.044***	−0.050***			−0.040***
						(0.003)	(0.004)			(0.004)

(*Continued*)

TABLE A11. (*Continued*)

	Baseline With Chapter Subsample	Moderator: Student–Teacher Ratio		Moderator: National Board Certification		Moderator: Teacher Experience		Moderator: Teacher Turnover		
		Main Effect	Interaction	Main Effect	Interaction	Main Effect	Interaction	Main Effect	Interaction	All Mods.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Teachers with <3 years experience × NC Pre-K funding							0.012** (0.004)			0.014*** (0.004)
Annual teacher turnover (in 10pp units)								−0.035*** (0.002)	−0.037*** (0.003)	−0.017*** (0.002)
Annual teacher turnover × NC Pre-K funding									0.006^{+} (0.004)	0.006 (0.003)
Observations	1,185,966	1,185,966	1,185,966	1,185,966	1,185,966	1,185,966	1,185,966	1,185,966	1,185,966	1,185,966
p Value (interactions = 0)										.001

Note. Standard errors are given in parentheses. Standard errors clustered at the county level. All models include Smart Start funding (in $000s), child and family covariates, county covariates, school-district covariates, and county, program year, and quarter fixed effects. See Table 3 for details. School district covariates: federal, state, and local per pupil expenditures in the school district.

NC Pre-K = North Carolina Pre-K.

$^{+}p < .10$.

$^{*}p < .05$.

$^{**}p < .01$.

$^{***}p < .001$.

TABLE A12

Effects of NC Pre-K Funding on Reading and Math Achievement: Moderation by County-Level Job Losses

	Baseline With Chapter Subsample	Moderator: County-Level Job Loss					
	(1)	Baseline (2)	Main Effect (3)	Interaction (4)	Baseline (5)	Main Effect (6)	Interaction (7)
Reading							
NC Pre-K funding ($000s)	0.052*	0.051+	0.051+	0.052*	0.005	0.005	0.007
	(0.026)	(0.027)	(0.027)	(0.026)	(0.023)	(0.023)	(0.023)
Job loss per full academic year (% affected)			0.000	0.000		0.002	0.001
			(0.003)	(0.003)		(0.003)	(0.003)
Percent affected by job loss per full academic year × NC Pre-K funding				0.005			0.015
				(0.016)			(0.012)
Observations	819,282	819,282	819,282	819,282	819,282	819,282	819,282
Math							
NC Pre-K funding ($000s)	0.059*	0.057*	0.057*	0.064*	0.008	0.008	0.013
	(0.027)	(0.027)	(0.027)	(0.026)	(0.024)	(0.024)	(0.024)
Job loss per full academic year (% affected)			0.006+	0.005		0.006+	0.004
			(0.003)	(0.003)		(0.004)	(0.004)
Percent affected by job loss per full academic year × NC Pre-K funding				0.027			0.039*
				(0.022)			(0.015)
Observations	823,297	823,297	823,297	823,297	823,297	823,297	823,297
Analysis specific controls							
Year of test FE		X	X	X	X	X	X
Time trends and quadratic					X	X	X

Note. Standard errors are given in parentheses. Standard errors clustered at the county level. All models include Smart Start funding (in $000s), child and family covariates, county covariates, and county, program year, and quarter fixed effects. See Table 3 for details.
NC Pre-K = North Carolina Pre-K.
+p < .10.
*p < .05.
**p < .01.
***p < .001.

TABLE A13

EFFECTS OF NC PRE-K FUNDING ON READING AND MATH ACHIEVEMENT: MODERATION BY COUNTY-LEVEL ECONOMIC FACTORS (IN THE YEAR OF THE TEST)

	Baseline With Chapter Subsample	Moderator: Median Family Income		Moderator: % SNAP Recipients		Moderator: % Medicaid Recipients		
		Main Effect	Interaction	Main Effect	Interaction	Main Effect	Interaction	All Mods.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Reading								
NC Pre-K funding ($000s)	0.034***	0.030**	0.018	0.046***	0.004	0.034**	0.031*	0.002
	(0.009)	(0.009)	(0.014)	(0.010)	(0.017)	(0.011)	(0.015)	(0.022)
Estimated median family income ($00s)		0.001***	0.001***					0.000^{+}
		(0.000)	(0.000)					(0.000)
Estimated median family income × NC Pre-K funding			−0.000					0.000
			(0.000)					(0.000)
SNAP recipients (%)				−0.016***	−0.020***			−0.018***
				(0.004)	(0.004)			(0.005)
SNAP recipients × NC Pre-K funding					0.004*			0.008**
					(0.002)			(0.003)
Medicaid recipients (%)						−0.023***	−0.023***	−0.008*
						(0.003)	(0.003)	(0.003)
Medicaid recipients × NC Pre-K funding							0.000	-0.004^{+}
							(0.001)	(0.002)
Observations	1,199,896	1,199,896	1,199,896	1,199,896	1,199,896	1,199,896	1,199,896	1,199,896
p Value (interactions = 0)								.015
Math								
NC Pre-K funding ($000s)	0.024*	0.022^{+}	−0.002	0.039**	−0.007	0.024^{+}	0.017	−0.016
	(0.012)	(0.012)	(0.019)	(0.014)	(0.021)	(0.014)	(0.019)	(0.026)

(*Continued*)

TABLE A13. (*Continued*)

	Baseline With Chapter Subsample	Moderator: Median Family Income		Moderator: % SNAP Recipients		Moderator: % Medicaid Recipients		
		Main Effect	Interaction	Main Effect	Interaction	Main Effect	Interaction	All Mods.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Estimated median family income ($00s)		0.001***	0.001***					0.000
		(0.000)	(0.000)					(0.000)
Estimated median family income × NC Pre-K funding			−0.000					−0.000
			(0.000)					(0.000)
SNAP recipients (%)				−0.017***	−0.021***			−0.019**
				(0.004)	(0.005)			(0.006)
SNAP recipients × NC Pre-K funding					0.004*			0.008*
					(0.002)			(0.003)
Medicaid recipients (%)						−0.023***	−0.023***	−0.008*
						(0.003)	(0.003)	(0.003)
Medicaid recipients × NC Pre-K funding							0.001	−0.005
							(0.001)	(0.003)
Observations	1,205,725	1,205,725	1,205,725	1,205,725	1,205,725	1,205,725	1,205,725	1,205,725
p Value (interactions = 0)								.039

Note. Standard errors are given in parentheses. Standard errors clustered at the county level. All models include Smart Start funding (in $000s), child and family covariates, county covariates, and county, program year, and quarter fixed effects. See Table 3 for details.
NC Pre-K = North Carolina Pre-K; SNAP = Supplemental Nutrition Assistance Program.
$^{+}p < .10$.
$^{*}p < .05$.
$^{**}p < .01$.
$^{***}p < .001$.

TABLE A14

Descriptive Statistics of Job Loss by Academic Year

Year	Mean (1)	*SD* (2)	Min (3)	Max (4)
1997–1998	0.575	0.498	0.000	1.910
1998–1999	0.813	0.819	0.000	3.546
1999–2000	0.868	0.857	0.000	5.174
2000–2001	1.426	1.099	0.000	5.929
2001–2002	1.426	1.099	0.000	5.929
2002–2003	1.143	0.914	0.000	8.137
2003–2004	0.894	0.991	0.000	5.659
2004–2005	0.657	0.645	0.000	4.160
2005–2006	0.832	0.689	0.000	4.903
2006–2007	0.730	0.760	0.000	7.060
2007–2008	0.617	0.596	0.000	6.772
2008–2009	1.349	0.887	0.000	4.892
2009–2010	0.703	0.446	0.000	2.689
2010–2011	0.733	0.557	0.000	3.737

Note. *SD* = standard deviation.

TABLE A15

EFFECTS OF NC PRE-K FUNDING ON ACADEMIC ACHIEVEMENT: MODERATION USING ALTERNATE MEASURES OF HEAD START (HS) AND EARLY HEAD START (EHS) SATURATION

	Moderator: Average EHS Enrollment		Moderator: 4-Year-Old HS Enrollment From PIR		Moderator: 4-Year-Old HS Enrollment Averaged From PIR and HS Cluster Methods		Moderator: 3-Year-Old HS Enrollment	
	Main Effect (1)	Interaction (2)	Main Effect (3)	Interaction (4)	Main Effect (5)	Interaction (6)	Main Effect (7)	Interaction (8)
NC Pre-K funding ($000s)	0.031**	0.031**	0.030**	0.029**	0.030**	0.038**	0.031**	0.033**
	(0.010)	(0.011)	(0.011)	(0.010)	(0.011)	(0.012)	(0.011)	(0.012)
Average EHS enrollment (in 10%)	−0.151***	−0.154***						
	(0.033)	(0.039)						
Average EHS enrollment × NC Pre-K funding		0.004						
		(0.030)						
4-year-old HS saturation (PIR) (in 10%)			0.103	−0.029				
			(0.265)	(0.216)				
4-year-old HS saturation (PIR) × NC Pre-K funding				0.223**				
				(0.084)				
Average of HS cluster and PIR calculations (in 10%)					−0.014	−0.007		
					(0.014)	(0.014)		
Average of HS cluster and PIR calculations × NC Pre-K funding						−0.024		
						(0.016)		

(Continued)

TABLE A15. (*Continued*)

	Moderator: Average EHS Enrollment		Moderator: 4-Year-Old HS Enrollment From PIR		Moderator: 4-Year-Old HS Enrollment Averaged From PIR and HS Cluster Methods		Moderator: 3-Year-Old HS Enrollment	
	Main Effect (1)	Interaction (2)	Main Effect (3)	Interaction (4)	Main Effect (5)	Interaction (6)	Main Effect (7)	Interaction (8)
3-year-old HS saturation (in 10%)							−0.006 (0.013)	−0.002 (0.013)
3-year-old HS saturation × NC Pre-K funding								−0.010 (0.010)
Observations	1,207,576	1,207,576	1,207,576	1,207,576	1,207,576	1,207,576	1,207,576	1,207,576

Note. Standard errors clustered at the county level. All models include Smart Start funding (in $000s), child and family covariates, county covariates, and county, program year, and quarter of birth fixed effects. See Table 3 for details.
NC Pre-K = North Carolina Pre-K.
$^{+}p < .10$.
$^{*}p < .05$.
$^{**}p < .01$.
$^{***}p < .001$.

TABLE A16

EFFECTS OF NC PRE-K FUNDING ON ACADEMIC ACHIEVEMENT: MODERATION BY ELEMENTARY SCHOOL CHARACTERISTICS WITH SCHOOL DEMOGRAPHIC CONTROLS

	Baseline With Chapter Subsample	Moderator: (Lagged) Academic Composite		Moderator: Local PPE		Moderator: State PPE		Moderator: Federal PPE		
		Main Effect	Interaction	Main Effect	Interaction	Main Effect	Interaction	Main Effect	Interaction	All Mods.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
NC Pre-K funding ($000s)	0.040**	0.023**	0.017*	0.040**	0.036**	0.040**	0.042**	0.040**	0.028*	0.001
	(0.012)	(0.008)	(0.008)	(0.012)	(0.011)	(0.012)	(0.012)	(0.012)	(0.012)	(0.009)
School-average achievement composite (lagged;std.)		0.218***	0.223***							0.222***
		(0.004)	(0.004)							(0.004)
School-average achievement composite (lagged;std.) × NC Pre-K funding			−0.010**							−0.004
			(0.003)							(0.003)
Local PPE in 000s ($2019)				0.031***	0.034***					−0.008^{+}
				(0.006)	(0.007)					(0.004)
Local PPE in 000s ($2019) × NC Pre-K funding					−0.006					0.003
					(0.009)					(0.008)
State PPE in 000s ($2019)						0.017*	0.019*			0.045***
						(0.008)	(0.008)			(0.008)
State PPE in 000s ($2019) × NC Pre-K funding							−0.003			−0.017**
							(0.005)			(0.006)

(*Continued*)

TABLE A16. (*Continued*)

	Baseline With Chapter Subsample	Moderator: (Lagged) Academic Composite		Moderator: Local PPE		Moderator: State PPE		Moderator: Federal PPE		
		Main Effect	Interaction	Main Effect	Interaction	Main Effect	Interaction	Main Effect	Interaction	All Mods.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Federal PPE in 000s ($2019)								−0.022	-0.036^{+}	−0.078***
								(0.019)	(0.021)	(0.018)
Federal PPE in 000s ($2019) × NC Pre-K funding									0.025	0.059***
									(0.017)	(0.015)
Observations	1,148,712	1,148,712	1,148,712	1,148,712	1,148,712	1,148,712	1,148,712	1,148,712	1,148,712	1,148,712
p Value (interactions = 0)										.003

Note. Standard errors are given in parentheses. Standard errors clustered at the county level. All models include Smart Start funding (in $000s), child and family covariates, county covariates, school-level covariates, and county, program year, and quarter fixed effects. See Table 3 for details. School-level demographic characteristics: % Black students, % Hispanic students, % economically disadvantaged students, average daily membership.
NC Pre-K = North Carolina Pre-K; PPE = per-pupil expenditure.
$^{+}p < .10$.
$^{*}p < .05$.
$^{**}p < .01$.
$^{***}p < .001$.

TABLE A17

EFFECTS OF NC PRE-K FUNDING ON ACADEMIC ACHIEVEMENT: MODERATION BY TEACHER CHARACTERISTICS WITH SCHOOL DEMOGRAPHIC CONTROLS

	Baseline With Chapter Subsample	Moderator: Student–Teacher Ratio		Moderator: National Board Certification		Moderator: Teacher Experience		Moderator: Teacher Turnover		
		Main Effect	Interaction	Main Effect	Interaction	Main Effect	Interaction	Main Effect	Interaction	All Mods.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
NC Pre-K funding ($000s)	0.040***	0.040***	0.040***	0.040***	0.040***	0.038***	0.037**	0.039**	0.038**	0.036**
	(0.012)	(0.012)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.012)	(0.012)	(0.011)
Student–teacher ratio in the school		−0.003*	-0.002^{+}							−0.001
		(0.001)	(0.001)							(0.001)
Student–teacher ratio in the school × NC Pre-K funding			-0.003^{+}							-0.003^{+}
			(0.002)							(0.002)
National Board Certified teachers (in 10pp units)				0.022***	0.019***					0.010*
				(0.003)	(0.005)					(0.005)
National Board Certified teachers × NC Pre-K funding					0.005					0.008
					(0.006)					(0.006)
Teachers with <3 years experience (in 10pp units)						−0.031***	−0.034***			−0.030***
						(0.003)	(0.003)			(0.003)

(*Continued*)

TABLE A17. (*Continued*)

	Baseline With Chapter Subsample	Moderator: Student–Teacher Ratio		Moderator: National Board Certification		Moderator: Teacher Experience		Moderator: Teacher Turnover		
		Main Effect	Interaction	Main Effect	Interaction	Main Effect	Interaction	Main Effect	Interaction	All Mods.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Teachers with <3 years experience × NC Pre-K funding							0.006 (0.004)			0.010** (0.003)
Annual teacher turnover (in 10pp units)								−0.021*** (0.003)	−0.019*** (0.003)	−0.009*** (0.002)
Annual teacher turnover × NC Pre-K funding									−0.004 (0.003)	−0.003 (0.003)
Observations	1,187,449	1,187,449	1,187,449	1,187,449	1,187,449	1,187,449	1,187,449	1,187,449	1,187,449	1,187,449
p Value (interactions = 0)										.000

Note. Standard errors are given in parentheses. Standard errors clustered at the county level. All models include Smart Start funding (in $000s), child and family covariates, county covariates, school-level covariates, and county, program year, and quarter fixed effects. See Table 3 for details. School-level demographic characteristics: % Black students, % Hispanic students, % economically disadvantaged students, average daily membership.

NC Pre-K = North Carolina Pre-K.

$^{+}p < .10$.

$^{*}p < .05$.

$^{**}p < .01$.

$^{***}p < .001$.

TABLE A18

Effects of NC Pre-K Funding on Academic Achievement: Moderation by School Achievement Averaged Across Grades 3–5

	Baseline With Chapter Subsample (1)	Moderator: School-Average Achievement Composite	
		Main Effect (2)	Interaction (3)
NC Pre-K funding ($000s)	0.030**	0.023**	0.016^{+}
	(0.011)	(0.008)	(0.008)
School-average achievement composite (lagged;std.;G3–G5 average)		0.212***	0.217***
		(0.004)	(0.004)
School-average achievement composite × NC Pre-K funding			−0.012***
			(0.003)
Observations	1,204,906	1,204,906	1,204,906

Note. Standard errors are given in parentheses. Standard errors clustered at the county level. All models include Smart Start funding (in $000s), child and family covariates, county covariates, and county, program year, and quarter fixed effects. See Table 3 for details.
NC Pre-K = North Carolina Pre-K.
$^{+}p < .10$.
$^{*}p < .05$.
$^{**}p < .01$.
$^{***}p < .001$.

TABLE A19

EFFECTS OF NC PRE-K FUNDING ON ACADEMIC ACHIEVEMENT: MODERATION BY COUNTY-LEVEL ECONOMIC FACTORS MEASURED IN CHILD'S YEAR OF BIRTH

	Baseline With Chapter Subsample	Moderator: Median Family Income		Moderator: % SNAP Recipients		Moderator: % Medicaid Recipients		
		Main Effect	Interaction	Main Effect	Interaction	Main Effect	Interaction	All Mods.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
NC Pre-K funding ($000s)	0.03040**	0.03040**	0.02065^{+}	0.03040**	0.02478*	0.03040**	0.02294	0.02609
	(0.01068)	(0.01068)	(0.01141)	(0.01068)	(0.01185)	(0.01068)	(0.01681)	(0.02316)
Estimated median family income ($00s)		−0.00029***	−0.00024***					−0.00023**
		(0.00007)	(0.00007)					(0.00008)
Estimated median family income × NC Pre-K funding			−0.00008					−0.00009
			(0.00005)					(0.00007)
SNAP recipients (%)				0.00556	0.00496			0.00584
				(0.00348)	(0.00343)			(0.00480)
SNAP recipients × NC Pre-K funding					0.00104			0.00140
					(0.00127)			(0.00439)
Medicaid recipients (%)						−0.00179	−0.00242	−0.00119
						(0.00337)	(0.00338)	(0.00420)
Medicaid recipients × NC Pre-K funding							0.00064	−0.00125
							(0.00114)	(0.00388)
Observations	1,207,576	1,207,576	1,207,576	1,207,576	1,207,576	1,207,576	1,207,576	1,207,576
p Value (interactions = 0)								.370

Note. Standard errors are given in parentheses. Standard errors clustered at the county level. All models include Smart Start funding (in $000s), child and family covariates, county covariates, and county, program year, and quarter fixed effects. See Table 3 for details. Due to the small effect sizes, we present coefficients with five digits following the decimal

NC Pre-K = North Carolina Pre-K; SNAP = Supplemental Nutrition Assistance Program.

$^{+}p < .10$.
$^{*}p < .05$.
$^{**}p < .01$.
$^{***}p < .001$.

TABLE A20

EFFECTS OF NC PRE-K FUNDING ON ACADEMIC ACHIEVEMENT: MODERATION BY COUNTY-LEVEL ECONOMIC FACTORS MEASURED IN CHILD'S YEAR OF PRE-K EXPOSURE

	Baseline With Chapter Subsample	Moderator: Median Family Income		Moderator: % SNAP Recipients		Moderator: % Medicaid Recipients		
		Main Effect	Interaction	Main Effect	Interaction	Main Effect	Interaction	All Mods.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
NC Pre-K funding ($000s)	0.03040**	0.03084**	0.01536	0.02570*	0.01363	0.03092**	0.02124	0.00782
	(0.01068)	(0.01010)	(0.01437)	(0.01023)	(0.01226)	(0.01068)	(0.01473)	(0.01393)
Estimated median family income ($00s)		0.00005	0.00003					0.00006
		(0.00013)	(0.00013)					(0.00014)
Estimated median family income × NC Pre-K funding			−0.00009					−0.00008
			(0.00007)					(0.00008)
SNAP recipients (%)				0.00759***	0.00647**			0.00742+
				(0.00223)	(0.00227)			(0.00389)
SNAP recipients × NC Pre-K funding					0.00150			0.00171
					(0.00128)			(0.00243)
Medicaid recipients (%)						0.00218	0.00203	0.00034
						(0.00248)	(0.00250)	(0.00439)
Medicaid recipients × NC Pre-K funding							0.00105	−0.00103
							(0.00120)	(0.00240)
Observations	1,207,576	1,207,576	1,207,576	1,207,576	1,207,576	1,207,576	1,207,576	1,207,576
p Value (interactions = 0)								.529

Note. Standard errors are given in parentheses. Standard errors clustered at the county level. All models include Smart Start funding (in $000s), child and family covariates, county covariates, and county, program year, and quarter fixed effects. See Table 3 for details. Due to the small effect sizes, we present coefficients with five digits following the decimal.

NC Pre-K = North Carolina Pre-K; SNAP = Supplemental Nutrition Assistance Program.

$^{+}p < .10$.
$^{*}p < .05$.
$^{**}p < .01$.
$^{***}p < .001$.

TABLE A21

EFFECTS OF NC-PRE-K FUNDING ON ACADEMIC ACHIEVEMENT OF BLACK STUDENTS BY AVERAGE ACHIEVEMENT OF THE SCHOOL

	All Black Students (1)	By Average Achievement of School: Below −1 *SD* (2)	Between −1 and 1 *SD* (3)	Above 1 *SD* (4)
NC Pre-K funding ($000s)	0.033*	0.019	0.042*	−0.057⁺
	(0.015)	(0.014)	(0.017)	(0.034)
Observations	349,373	97,657	208,351	43,361

Note. Standard errors are given in parentheses. Standard errors clustered at the county level. All models include Smart Start funding (in $000s), child and family covariates, county covariates, and county, program year, and quarter fixed effects. See Table 3 for details.
NC Pre-K = North Carolina Pre-K; *SD* = standard deviation.
$^{+}p < .10$.
$^{*}p < .05$.
$^{**}p < .01$.
$^{***}p < .001$.

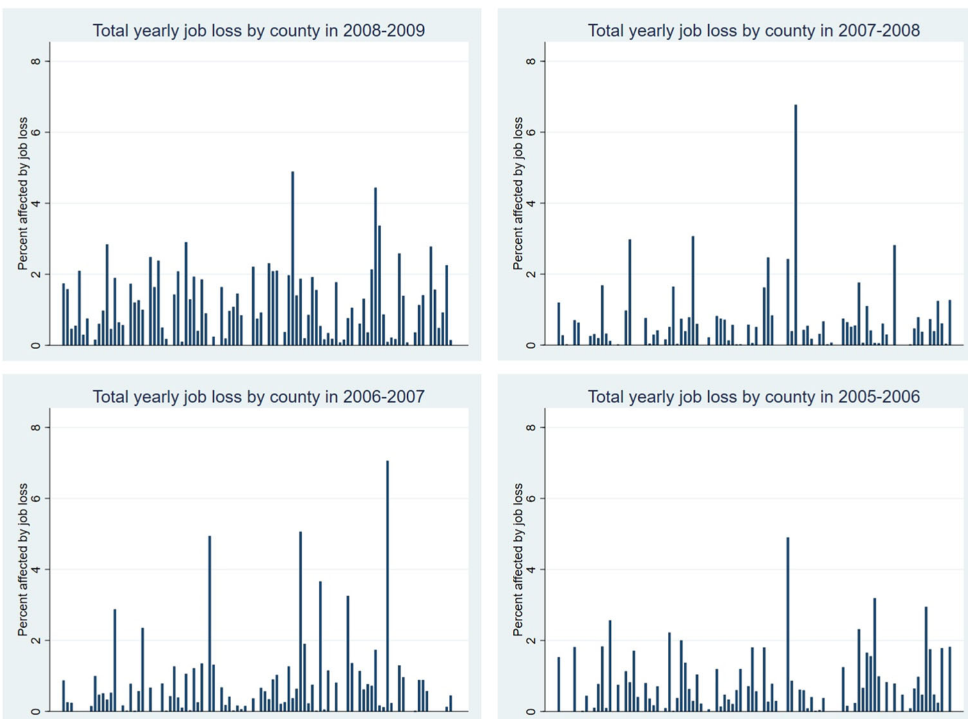


FIGURE A1.—Job loss across four years in all counties in North Carolina.

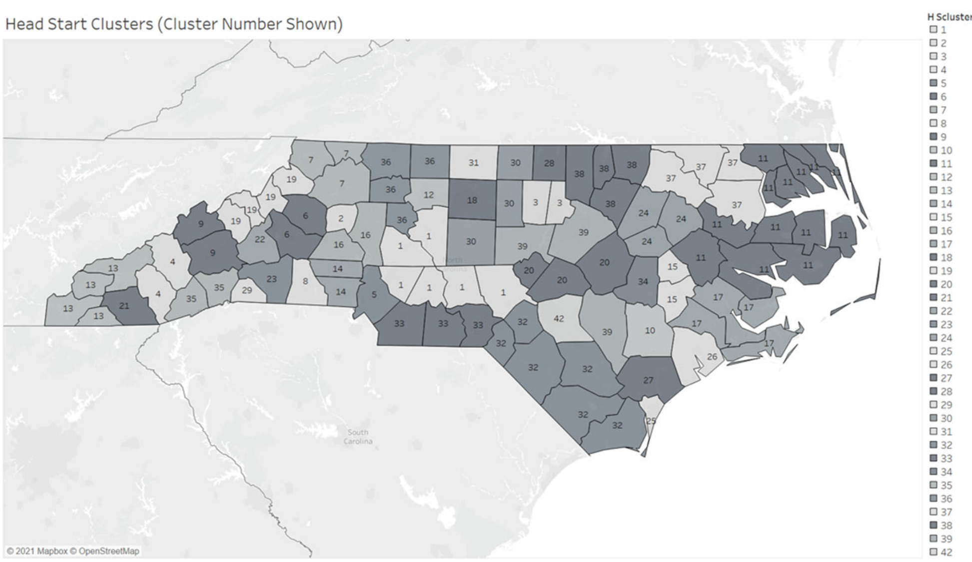
APPENDIX B: ADDITIONAL DETAILS FOR THE COUNTY-LEVEL HEAD START AND EARLY HEAD START MEASURES

A. Creating Head Start Service Clusters Based on Center and Grantee-level Data

We provide here an overview of our data processing to create the Head Start (HS) and Early Head Start (EHS) service region county groups, which we term clusters. We then explain how we calculated HS and EHS enrollment saturation for these clusters, and then disaggregated these modified calculations into counties for the moderation analyses.

Step 1. Identifying HS/EHS county clusters (separately for HS/EHS)

1. Download CSV file from HS/EHS program map (June 2021)
2. Subset to NC program sites
3. Sort by HS/EHS grantee
4. Identify the service region (i.e., counties) served by each HS/EHS grantee (i.e., the counties where a HS/EHS grantee was listed as operating a HS/EHS program site)
5. Call all HS/EHS grantees to confirm their service regions, focus on contacting HS/EHS grantees where we identified overlap in their service regions. Edit as needed
6. Code the *HScluster* variable to identify a grantee as part of an HS/EHS service regions (illustrated in figure below)
7. Code the *HSunique* variable to identify HS/EHS grantees that served only the county where their administrative office was located, even if that county was also served by another HS/EHS grantee



Step 2. Converting Grantee-level HS data into County-level HS Data

1. Merge PIR data file and the HScluster file by county FIPS code
2. Create separate files for counties by their status of HSunique (one data set for HSunique = 1 and another for HSunique = 0)
3. For counties with HSunique = 0, disaggregate 3-year-old HS enrollment numbers across counties within each cluster based on the % of 3-year-old children in each county within the cluster (repeat 4-year-old HS enrollment)
4. After disaggregating 3-year-old HS enrollment #s for HSunique = 0 cases across counties within clusters, add enrollment #s for HSunique = 1 cases to the sum of enrollment #s for each relevant county (repeat 4-year-old HS enrollment)
5. Divide county-level 3-year-old HS enrollment #s by the # of 3-year-old children in each county to calculate a HS penetration variable (repeat 4-year-old HS penetration)

B. Annual Descriptive Statistics for Head Start and Early Head Start Enrollment Measures Derived From the Cluster-Disaggregation Process and Compared With Raw (Unadjusted) PIR Enrollment

TABLE B1
Descriptive Statistics: County-Level Head Start (HS) Exposure

Program Year	Age-3 HS Enrollment With Clusters *M*	Age-3 HS Enrollment With Clusters (*SD*)	Age-4 HS Enrollment With Clusters *M*	Age-4 HS Enrollment With Clusters (*SD*)	Age-4 HS Enrollment From PIR (Raw) *M*	Age-4 HS Enrollment From PIR (Raw) (*SD*)
1993	0.05	(0.05)	0.12	(0.06)	0.08	(0.16)
1994	0.04	(0.05)	0.12	(0.06)	0.10	(0.17)
1995	0.04	(0.06)	0.12	(0.07)	0.10	(0.18)
1996	0.05	(0.06)	0.13	(0.13)	0.12	(0.30)
1997	0.08	(0.08)	0.13	(0.06)	0.12	(0.23)
1998	0.06	(0.06)	0.12	(0.07)	0.11	(0.20)
1999	0.06	(0.05)	0.12	(0.07)	0.11	(0.19)
2000	0.07	(0.06)	0.12	(0.06)	0.11	(0.17)
2001	0.07	(0.06)	0.14	(0.10)	0.12	(0.19)
2002	0.08	(0.05)	0.14	(0.07)	0.12	(0.19)
2003	0.09	(0.06)	0.15	(0.07)	0.12	(0.20)
2004	0.09	(0.07)	0.14	(0.06)	0.11	(0.19)
2005	0.08	(0.06)	0.14	(0.06)	0.11	(0.20)
2006	0.08	(0.05)	0.14	(0.06)	0.11	(0.20)
2007	0.08	(0.05)	0.13	(0.06)	0.11	(0.19)
2008	0.08	(0.05)	0.13	(0.07)	0.10	(0.20)
2009	0.09	(0.06)	0.12	(0.06)	0.10	(0.19)
2010	0.09	(0.05)	0.13	(0.06)	0.10	(0.16)

Note. $N = 100$ counties per program year. Program year corresponds to the year in which children were age 4. Cumulative HS exposure is the average of age-4 and age-3 HS exposure. Values displayed for age-3 HS exposure were lagged from the age-3 program year.
SD = standard deviation.

TABLE B2

Descriptive Statistics: County-Level Early Head Start (EHS) Exposure

Program Year	Maximum EHS Exposure		Cumulative EHS Exposure		Age-2 EHS Exposure		Age-1 EHS Exposure		Age-0 EHS Exposure	
	M	(*SD*)	*M*	(*SD*)	*M*	(*SD*)	*M*	(*SD*)	*M*	(*SD*)
1993	0.00	(0.00)	0.00	(0.00)	0.00	(0.00)	0.00	(0.00)	0.00	(0.00)
1994	0.00	(0.00)	0.00	(0.00)	0.00	(0.00)	0.00	(0.00)	0.00	(0.00)
1995	0.00	(0.00)	0.00	(0.00)	0.00	(0.00)	0.00	(0.00)	0.00	(0.00)
1996	0.00	(0.00)	0.00	(0.00)	0.00	(0.00)	0.00	(0.00)	0.00	(0.00)
1997	0.00	(0.01)	0.00	(0.00)	0.00	(0.01)	0.00	(0.00)	0.00	(0.00)
1998	0.01	(0.05)	0.00	(0.02)	0.01	(0.05)	0.00	(0.01)	0.00	(0.00)
1999	0.00	(0.02)	0.00	(0.01)	0.00	(0.01)	0.00	(0.02)	0.00	(0.01)
2000	0.00	(0.02)	0.00	(0.01)	0.00	(0.01)	0.00	(0.01)	0.00	(0.02)
2001	0.00	(0.02)	0.00	(0.01)	0.00	(0.01)	0.00	(0.01)	0.00	(0.02)
2002	0.00	(0.03)	0.00	(0.02)	0.00	(0.03)	0.00	(0.01)	0.00	(0.01)
2003	0.00	(0.02)	0.00	(0.02)	0.00	(0.02)	0.00	(0.02)	0.00	(0.01)
2004	0.01	(0.03)	0.00	(0.02)	0.01	(0.03)	0.00	(0.02)	0.00	(0.02)
2005	0.01	(0.02)	0.00	(0.02)	0.01	(0.02)	0.00	(0.02)	0.00	(0.01)
2006	0.01	(0.02)	0.00	(0.02)	0.01	(0.02)	0.00	(0.02)	0.00	(0.01)
2007	0.01	(0.03)	0.01	(0.02)	0.01	(0.03)	0.01	(0.02)	0.00	(0.01)
2008	0.01	(0.02)	0.00	(0.01)	0.01	(0.01)	0.01	(0.02)	0.00	(0.01)
2009	0.01	(0.03)	0.01	(0.02)	0.01	(0.02)	0.01	(0.02)	0.00	(0.01)
2010	0.01	(0.03)	0.01	(0.02)	0.01	(0.03)	0.01	(0.03)	0.00	(0.01)

Note. $N = 100$ counties per program year. Program year corresponds to the year in which children were age 4. Maximum EHS exposure is the maximum value recorded across age-2, age-1, and age-0 EHS exposure. Cumulative EHS exposure is the average of age-2, age-1, and age-0 EHS exposure. Values displayed for age-2, age-1, and age-0 EHS exposure were lagged from the age-2, age 1, and age-0 program years, respectively.
SD = standard deviation.