HHS Public Access

Author manuscript

Child Youth Serv Rev. Author manuscript; available in PMC 2023 July 20.

Published in final edited form as:

Child Youth Serv Rev. 2020 May ; 112: . doi:10.1016/j.childyouth.2020.104890.

Long-term effects of early childhood programs through eighth grade: Do the effects fade out or grow?

Yu Bai,

Helen F. Ladd,

Clara G. Muschkin,

Kenneth A. Dodge*

Duke University, Sanford School of Public Policy, 201 Science Dr, Durham, NC 27708, United States

Abstract

Support for policies to improve early childhood educational development and reduce disparities grew rapidly this century but recently has wavered because of findings that program effects might fade out prematurely. Two programs implemented at scale in North Carolina (Smart Start and More at Four) have been associated with academic success early in elementary school, but it is not known whether these effects fade out or are sustained in middle school. Smart Start provides state funding to support high-quality early childcare in local communities, and More at Four provides state-funded slots for a year of credentialed pre-kindergarten. Funds were allocated for each program at varying rates across counties and years. We used this variation to estimate the long-term impact of each program through eighth grade, by measuring the association between state funding allocations to each program, in each of 100 counties over each of 13 consecutive years, and later student performance. Students were matched to funding levels provided to their home county in their early childhood years and then followed through eighth grade. Analyses using county- and year-fixed-effects regression models with individualand school-level covariates conducted on nearly 900,000 middle school students indicate significant positive impacts of funding for each program on reading and math test scores and reductions in special education placement and grade retention. These impacts do not fade out and seem instead to grow (for More at Four) as students progress through middle school. Students from economically disadvantaged backgrounds experience particularly large benefits from the More at Four Program.

*Corresponding author: dodge@duke.edu (K.A. Dodge).

CRediT authorship contribution statement

Yu Bai: Data curation, Formal analysis, Methodology, Writing original draft, Writing - review & editing. **Helen F. Ladd**: Conceptualization, Methodology, Investigation, Resources, Writing review & editing, Supervision. **Clara G. Muschkin**: Conceptualization, Methodology, Investigation, Resources, Writing - review & editing, Supervision. **Kenneth A. Dodge**: Conceptualization, Methodology, Investigation, Resources, Writing - review & editing, Supervision.

Declaration of Competing Interest

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

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.childyouth.2020.104890.

Keywords

Early childhood; Smart Start; More at Four; Middle school; Academic success

1. Introduction

Over the past several decades, early child care and education policies have provided increased support for programs to promote the cognitive and social-emotional development of young children from birth to the time they enter kindergarten. Such policies reflect the growing knowledge of the importance of environmental experiences in nurturing early brain development for later school success (Follari, 2015; Stern, 2017). Following encouraging findings from small-scale experiments by private funders, numerous communities and states have scaled up programs with public resources such as preschool programs (U.S. Government Accountability Office, 2019). Most evaluations of scaled-up early childhood programs demonstrate positive impact at the end of the early intervention period, with promise to close achievement gaps across income and ethnic groups (Dodge, 2017). However, disappointment ensued when the impacts of some programs, particularly Head Start (U.S. Department of Health and Human Services, 2010) and Tennessee's Voluntary Pre-K Program (Lipsey, Farran, & Durkin, 2018), were found to fade out after participants enter elementary school. Due to funding issues, methodological challenges, and the recency of some studies, very few evaluations have followed participants past the primary grades, calling into question whether even the most "successful" early childhood programs have long-lasting effects. The current study evaluated the impact of two scaled-up, statewide early childhood programs on outcomes through the end of middle school.

1.1. More at Four and Smart Start

Two early childhood intervention programs have been pioneered and scaled up in North Carolina: Smart Start (SS) was initiated in 1993, and More at Four (MF, since called NC pre-K) started in 2001. The programs cover rural, suburban, and urban areas. Started in 18 counties in 1993, SS spread to more than 50 counties by 1997, and covered all 100 NC counties by 1998. SS allocated state dollars to entities that were administered at the county-level annually to improve the quality and availability of early childcare services for children from birth to age 4 in order to ensure that all children enter school healthy and ready to learn (Ladd, Muschkin, & Dodge, 2014). The SS program spent 52 million dollars in child care related activities in FY 2017–2018 (North Carolina General Assembly, 2019). It served 1974 child care programs to improve the quality of care, provided 24,427 children with subsidy assistance, and offered 4732 teachers an education-based salary supplement.

MF was established in 2001 and is designed to provide a highquality education experience to promote school readiness for disadvantaged four-year-old children living in North Carolina. MF provides state-funded slots for year-long pre-kindergarten for high-risk 4-year-old children as long as the pre-k classroom meets high standards of quality (Peisner-Feinberg & Schaaf, 2006). Many pre-k classrooms readily sought these funded slots by improving their standards in class size, teacher credentials, and curriculum. Most qualifying pre-k classrooms also enrolled children who were not subsidized by the state; thus, the MF

program directly affected children who received funded slots and additional children who were enrolled in the same classrooms. North Carolina preschool enrolled 28,385 children in 2017–2018, about a quarter of the population of four-year-olds statewide (Friedman-Krauss et al., 2019). The MF program costed about $5300 per participating child and enrolled 22 percent of 4-year-olds in 2016 (Barnett et al., 2017). We note for comparison that in 2016, Head Start cost between $8000 and $10,000 per participating child (Congressional Budget Office, 2017) and Early Head Start cost $12,757 per child (Friedman-Krauss, 2016). Only about 10% of the eligible population actually participated in the Head Start programs in 2016.

1.2. Theoretical framework

Our framework is based on a dynamic model of skill formation (Cunha & Heckman, 2007; Cunha & Heckman, 2008). A child's skill is acquired through a multistage process. Inputs at each stage produce outputs at the next stage. Some skills may be more productively learned at some stages than other stages, and some investments may be more productive at some stages than at other stages. Skills acquired in one stage can persevere and afford learning of new skills at future stages Cunha and Heckman (2007,2008) call this perspective "skill begets skill" because basic skills are required to learn more advanced skills, such that trajectories of differences between initial-skill learners and non-learners will diverge over time. Initial intervention effects will not fade out but rather will grow.

The skill formation theory posits the lasting significance of early intervention. For example, early language and literacy skills are an important indicator of school readiness (Schickedanz, 1999). These skills develop in early years of life, and development is a continuous formation process (Schickedanz, 1999). Infants learn how to turn pages. Toddlers learn how to pay attention and interact with pictures in books. Preschoolers develop skills of story comprehension by re-telling stories, and school-aged children learn to read. Children who have access to early experiences and environments that foster and promote their early language and literacy development will outperform other children not only at school entry but across the lifespan.

In contrast, Bailey, Duncan, Odgers, and Yu (2017) suggest circumstances in which fade out is expected to occur. If the skill that is taught early is one that almost all children will learn soon anyway, such as walking, then early intervention might help some children precociously acquire a skill, but when other children acquire the skill naturally the differences between groups will converge. Similarly, children's environments might be naturally structured, through evolution, to support those children who lag behind their peers by giving them special attention, so that they "catch up" and all children reach a minimal level of competence.

The Smart Start and More at Four programs are designed to build foundational skills. Smart Start improves the quality of early childcare services for children from birth to age 4. More at Four provides qualified teachers, structured curricula, and formal instruction for children at age 4. Both programs are designed to ensure that all children entering kindergarten will be ready to learn. Past evaluations of SS and MF show that children exposed to higher levels of funding for each program display more advanced reading and math achievement in

elementary school (Dodge, Bai, Ladd, & Muschkin, 2017). Because skills developed in one period can persist and afford greater learning at future periods (Cunha & Heckman, 2007; Cunha & Heckman, 2008), we hypothesize persistent positive program effects for both SS and MF in their middle-school education outcomes.

1.3. Literature review

The long-term positive impact of much-publicized early interventions, such as the Perry Preschool Project (Heckman, Moon, Pinto, Savelyev, & Yvitz, 2010; Schweinhart et al., 2005), Abecedarian (Campbell et al., 2014), and the Child-Parent Center (CPC) program (Reynolds, Ou, Mondi, & Giovanelli, 2019) gives support to the skills-beget-skills hypothesis. These programs provide proof-of-concept for the merit of early interventions producing long-term impact. However, evidence from these interventions may have limited relevance for current policy because their evaluations were conducted four decades ago by private agencies with highly disadvantaged, stay-at-home mothers at very high funding levels per child and comparatively small samples. Although they have inspired the design of programs that are being implemented at scale today, none of the scaled-up programs are directly comparable to the small and expensive early programs. Today, programs are implemented publicly at scale with different curricula and more modest budgets per child, requiring new evaluation.

1.3.1. More at Four and other pre-kindergarten programs—In previous publications, we evaluated the impact of MF on children's educational attainments as these students progressed through elementary school by estimating the impact of the allocation of state funds for MF to a county in a given year on eligible student educational outcomes (Dodge et al., 2017; Ladd et al., 2014; Muschkin, Ladd, & Dodge, 2015; Muschkin, Ladd, Dodge, & Bai, 2018). We found that a greater allocation of state funds to a county in a year was associated with better outcomes for children in that county-year as they progressed through elementary school in grades three, four, and five. A larger allocation was associated with better, end-of-grade (EOG) standardized test scores in reading and math and reductions in the probability of grade retention and placement into special education services.

The positive effects of the MF program conducted in NC have been confirmed by other researchers. Using a matched comparison design, Peisner-Feinberg et al. (Peisner-Feinberg et al., 2014; Peisner-Feinberg, Garwood, & Mokrova, 2016; Peisner-Feinberg, Mokrova, & Anderson, 2017) found that MF participants had better language, literacy, and math skills in kindergarten entry evaluation than those who did not participate in the program, and these effects persisted through the end of kindergarten (Peisner-Feinberg, Schaaf, Hildebrandt, Pan, & Warnaar, 2015) and through third grade as measured by math and reading scores on standardized end-of-grade tests (Peisner-Feinberg & Schaaf, 2010). However, the research on MF effects has not yet been extended to middle school.

State-administered pre-kindergarten programs have been operated in other states for a long time, but findings of long-term program effects have been mixed. Positive impacts in elementary school have been observed in Arkansas (Jung, Barnett, Hustedt, & Francis, 2013), Colorado (Atteberry, Bassok, & EWiong, 2019), Delaware (GamelMcCormick,

Amsden, & Hartranft, 2005), Georgia (Early, Li, & Maxwell, 2017), Louisiana (Ramey, Ramey & Asmus, 2011), New Jersey (Barnett, Jung, Youn, & Frede, 2013), and Texas (Andrews, Jargowsky, & Kuhne, 2012). Moreover, those effects did not disappear, and extended to middle school students in Nevada (Leitner, 2013) and Oklahoma (Gormley, Phillips, & Anderson, 2018). The positive effects were observed even in high school students in Colorado (Atteberry, Bassok, & EWiong, 2019) and Michigan (Schweinhart, Xiang, Daniel-Echols, Browning, & Wakabayashi, 2012). The long-term effects of the Great Start Readiness Program (GSRP), a state-funded preschool program sponsored by the state of Michigan, were examined through high school graduation (Schweinhart et al., 2012). Compared to nonparticipants, GSRP participants were more likely to graduate from high school on time, less likely to experience grade retention, and have a higher level of proficiency on the Michigan Merit Examination. Students who participated in the Colorado Preschool Program were less likely to be identified as reading deficient during kindergarten and to experience grade retention during K-3, and more likely to graduate from high school on time than their nonparticipant peers (Atteberry, Bassok, & EWiong, 2019). The pre-kindergarten program in Nevada was found to have positive effects that extended to seventh grade (Leitner 2013). Gormley et al. (2018) found enduring positive effects of Tulsa's pre-kindergarten program on math performance, enrollment in honors courses, and reductions in grade retention.

In contrast, other studies have shown that positive effects of prekindergarten programs soon faded. A randomized control trial of the Tennessee Voluntary Pre-K Program (VPK) (Lipsey et al., 2018) found that VPK students performed better in literacy, language, and math during the pre-k year than control students, but those positive VPK effects disappeared by the end of kindergarten. Moreover, the academic performance advantage between the two groups reversed by second grade and third grade. A study of Florida's Voluntary Pre-kindergarten Program showed that the program failed to decrease the likelihood of grade retention during the period between kindergarten and third grade (Miller & Bassok, 2019). No differences in academic performance were found between third graders who had attended Iowa's pre-kindergarten program and those who had not (American Institutes for Research, 2013).

1.3.2. Smart Start and other early childhood programs—Smart Start serves children birth to four with the goals of raising the quality of childcare, strengthening families, advancing child health and development, and improving early literacy (North Carolina Partnership for Children, 2019). Over the time period of implementation of Smart Start, the quality of childcare has improved (Bryant, Bernier, Peisner-Feinberg, & Maxwell, 2002). The percentage of children in highestrated childcare programs (4 or 5 stars on a 5-star scale) increased from 33 percent in 2001 to 72 percent in 2019 (North Carolina Partnership for Children, 2019). The average star rating for child care facilities sponsored by SS increased from 2.68 in 2001 to 4.52 stars in 2019 (North Carolina Partnership for Children, 2019).

We have evaluated the effect of SS on children's educational attainments with the same design that we used to evaluate the effects of MF, by estimating the association between the allocation of state funds for SS to a county in a given year with student educational outcomes (Dodge et al., 2017; Ladd et al., 2014; Muschkin et al., 2015; Muschkin et al.,

2018). We found that a greater allocation of state funds to a county in a year was associated with better end-of-grade (EOG) standardized test scores in reading and math and reductions in the probability of grade retention and placement into special education services, in grades 3,4, and 5.

Other research evaluating SS effects on elementary-school education outcomes is consistent with these findings. Maxwell and her colleagues examined the effects of SS on kindergarten entry skills among children living in six counties of North Carolina (Maxwell, Bryant, & Miller-Johnson, 1999). They found that kindergarteners who had participated in SS had significantly better cognitive and language skills than their comparison peers. Moreover, fewer children in the SS group were rated to have behavior problems by classroom teachers than children in the comparison group. Another study evaluated the associations between children's academic performance and the quality of SS-sponsored childcare centers (Bryant et al., 2003). The authors found that children in higher quality centers scored significantly higher on school-readiness measures than children from lower quality centers. These findings support the positive effects of SS on children's outcomes. However, these evaluations focused only on the short term and did not examine outcomes beyond elementary school.

1.3.3. Early childhood program for disadvantaged children—

Little research has examined the effects of pre-K enrollment among disadvantaged students as they progress through middle school. Gormley et al. (2018) reported gains from Tulsa's pre-k program among middle-school students with socioeconomic disadvantage. Pre-K enrollment improved math test scores and reduced likelihood of grade retention among students with free/reduced lunch. Similar findings were reported for English language learners and minority students.

1.4. Research hypotheses

One hypothesis guiding policy has been that early childhood educational experiences propel a trajectory of increasing positive impact by building skills that, in turn, build more skills over the life course (Heckman, 2006; Shonkoff & Phillips, 2000). A counter-hypothesis posits that those skills would eventually develop even without structured intervention, such that intervention effects fade out when other children "catch up" on their own. Another counter-hypothesis is that later school environments tend to subvert positive effects of early intervention in unspecified ways, perhaps by neglecting students who come to school ready to learn and favoring remediation of ill-prepared students. Still another hypothesis is that the nature and quality of early childhood education programs differ across states, leading to accurate, but inconsistent, findings across evaluations because the programs themselves differ (Dodge, 2017).

The current study tests the persistence, acceleration, and fadeout hypotheses by extending the evaluation of MF and SS over the course of middle school through the end of grade 8. We test program impact on math and reading test scores, the likelihood of being placed in special education services, and the probability of being a grade repeater. We compared effect sizes across grades 3–8 to observe whether impacts remain stable (persistence), grow

(acceleration), or diminish (fadeout). We also asked whether program impacts differ across subgroups within the population, defined by maternal education attainment level, race, and family income, in order to understand whether these early childhood programs contribute to reducing achievement gaps.

2. Method

2.1. Study population

The study population is composed of all sixth, seventh, and eighth graders who had been born in North Carolina between January 1, 1988, and December 31, 2000, and who also enrolled in a North Carolina public school (including charter schools) any time between school year 1999–2000 and school year 2015–2016. Identified birth records and school records of individual children were sent to the North Carolina Education Research Data Center, where they were matched, de-identified, and provided to the researchers. All procedures were pre-approved by the Duke University IRB and the State of North Carolina Department of Public Instruction. Previously, we have reported that about 74% of birth records could be matched to their elementary school records, a figure that approaches the expected maximum because of expected attrition related to moving out-of-state or enrolling in private schools. In the current study, 76% of birth records were matched to middle school records (slightly higher, perhaps because some students re-entered public school in middle grades). About 900,000 students are included in the study panel for each grade: 907,738 in Grade 6, 902,865 in Grade 7, and 896,349 in Grade 8.

2.2. Independent variables

A student's exposures to each program were measured by separate state allocations per-child for SS and MF, respectively, to a county in the year(s) that a student qualified to benefit from allocations. These variables were matched to a student based on the age or ages at which the student was eligible for the specified program based on the student's county of birth (five years across ages 0–4 for SS and one year at age 4 for MF).

SS funding varied across years from zero to a peak of $250 million (in 2009$) in 2000. The average funding per child was $220 per child aged 0 to 4 per year in 2009 (the last year of program funding included in the current evaluation). Because funding could cover a child for each of five years across ages 0–4, the total state funds allocated per child within a county averaged $1100($220 × 5 years) and ranged from $0 to $3500, with high variance within and between counties over the study period.

Like SS, MF started in selected pilot counties and grew across years to all 100 counties so that the number of counties served and dollars allocated per student varied across counties and years. Funding for an MF slot for an eligible four-year-old child to attend a certified prekindergarten program averaged $4400 for one year. The program reached about 25% of all four-year-olds by 2010; therefore, the average MF investment during years of this evaluation was about $1100 ($4400 × 25%), similar to the SS investment. We used $1100 per child for both SS and MF when we calculated the effect sizes of the state's investment. All analyses use the continuous variable in dollars.

2.3. Dependent variables and covariates

Covariates include individual student characteristics recorded annually (gender, race/ethnicity, and economically disadvantaged status as designated by eligibility for free/reduced lunch), individual mother characteristics documented in vital records when the child was born (infant birth weight, years of mother's education, marital status, age, information of infant's birth father, immigration status, first born status, and mother's racial group), school-level characteristics (percen of non-Hispanic Black students, percent of Hispanic students, and charter school status), and birth-county characteristics (percent of births to Black mothers, percent of births to Hispanic mothers, percent of births to low education mothers, number of births, total population, median family income, percent of population receiving food stamps, and percent of population receiving Medicaid).

Dependent variables were end-of-grade standardized test scores in reading and math in grades 6,7, and 8; special education placement in each grade and ever since Grade 3; and grade retention in each grade and ever since Grade 3.

2.4. Analysis plan

Analytic models were estimated as linear regression models when the dependent variables are reading and math standardized test scores and as logistic regression models for dichotomous outcome variables. Independent variables are the SS and MF allocation amount, calibrated to 2009 dollars. Additional details about the variables and models have been described previously (Dodge et al., 2017; Ladd et al., 2014; Muschkin et al., 2015; Muschkin et al., 2018). We applied listwise deletion of missing data because only 0.7% of observations had one or more missing values in control variables.

We tested impact of SS and MF in a model that takes the following form,

$$O_{icbtg} = \beta_1 SS_{ict}{}^{*} + \beta_2 MF_{ict}{}^{*} + \beta_3 X_{ib} + \beta_4 Y_{it} + \beta_5 C_{cb} + Grade_{ig} + \alpha_c + \gamma_b + \varepsilon_{icbt}.$$

where O_{icbtg} is a grade g outcome in year t for the i^{th} student born in county c in year b. SS_{ict} * is a Smart Start variable for the i^{th} child, MF_{ict} * is a More at Four variable. X_{ib} is a vector of characteristics of the i^{th} child at the time of birth including, for example, the child's birth weight and the education level of the child's mother. Y_{it} are characteristics of the i^{th} child observed in year t, such as the race of the child as reported on school records. All models included county fixed effects (α_c) and year fixed effects (γ_b) in order to control for any overall county differences and secular trends in outcomes. The analyses thus provide tests of differences in SS and MF funding within counties across time, netting out any statewide time trends.

All models included fixed covariates at the student level (X, i.e., child gender (reference = Female); birth weight as extremely, very low, low, or high (reference = normal birth weight); race as black, Hispanic, Asian, American Indian, or other (reference = White); economically disadvantaged (1 = yes); mother's education level as continuous variable; marital status (1 = Married); age as continuous variable; father information available on birth record (1 = Yes); immigration status (1 = Yes); first born status (1 = Yes); and

race), and time-varying covariates at the county level C, i.e., proportion births to black mothers, proportion births to Hispanic mothers, proportion births to low education mothers, number of births, total population, median family income, population with food stamps, and population with Medicaid), in order to control for any factors that might correlate with funding allocations.

We examine subgroup differences in program effects by including interaction terms for education levels of the mother, mother's racial group, and student's family economically disadvantaged status (indicated by qualifying for free or reduced-price lunch).

Just as in previous research, we conducted sensitivity tests to determine whether the findings hold under alternate assumptions about program exposure. About 21.2% of children moved from one NC county to another between birth and grade 3 (Dodge et al., 2017; Muschkin et al., 2015). We conducted one set of analyses assuming that a student lived in the county of birth when exposed to each program, and a second set of analyses assuming that the child lived in the county of school attendance at time of exposure to each program.

3. Results

3.1. Estimated associations between program exposure and academic outcomes, by grade level

Table 1 summarizes the descriptive information on academic outcomes by grade. Table 2 lists means, standard deviations, and sample sizes for all other variables. We find that increases in per-child funding for each program (SS and MF) are associated with higher student test scores in reading and math in each of Grades 6,7, and 8 (each $p < 0.01$, Table 3). Estimated coefficients are similar or increase across the three middle school years, and the magnitude of effects is at least as large as in elementary school (see Dodge et al., 2017), indicating no fadeout of impact across middle school years.

We illustrate the cross-age effect sizes in Fig. 1. Because the average gain in test scores in one school year is about 0.5 standard deviations in reading and 0.75 standard deviations in math (Ladd et al., 2014) and the school year lasts 10 months, an effect size of 0.1 standard deviations represents about 2 months of learning in reading and 1.5 months in math. Fig. 1 depicts the effect sizes in months of added learning associated with the average funding allocation (as compared with zero funds) allocated for SS and MF separately, for each of reading and math scores at Grades 3–8, with figures for Grades 3–5 taken from previous publications (Dodge et al., 2017). One observes no fading of effect across Grades 3–8; in contrast, effects for MF appreciably grow.

MF funding allocation was associated with reduced probability of repeating eighth grade $p < 0.01$ by 5.6%, but there were no effects on grade retention for either sixth or seventh grades. Each program was associated with a reduced probability of ever being grade retained at least one time between Grade 3 and Grade 8 (each $p < 0.01$), with magnitudes of 3.8% to 10.1% (Table 3).

Each program allocation was associated with a reduced probability of being placed into special education services in Grades 6, 7, and 8 (each $p < 0.01$) by up to 5.0% and a decreased probability of ever receiving special education services between Grade 3 and Grade 8 (each $p < 0.01$) by up to 4.1% (Table 3).

3.2. Program effects for sub-populations

To examine whether SS and MF funding levels had different associations for subgroups defined by education levels of the mother (less than high school graduation = low, otherwise high), mother's racial group (Black vs. not), or student's family economic status (free or reduced-price lunch status vs. not), we add group interaction terms to the basic models (six terms in all: SS [and separately MF] × mother's education, mother's race, and student's economic status). Two consistent statistically significant interaction patterns emerged for MF Students who were economically disadvantaged gained more from exposure to MF in math and reading scores in grades 6, 7, and 8 than their peers who were not economically disadvantaged (each $p < 0.01$, Table 4). Effects for economically advantaged students were also positive but not as large. Fig. 2 depicts this interaction for reading scores.

Compared to their peers, those students whose mothers were Black exhibited larger positive impacts from MF on math and reading scores at grades 6,7, and 8 (each $p < 0.01$). Effects for non-Black student were also positive but not as large.

No interaction term was significant for SS, indicating that program effects did not differ significantly across subgroups.

3.3. Sensitivity analysis

Nearly one fourth of students in the grade panels changed their family residence from their county of birth to their county of school attendance (Table 2). In reported analyses, we coded a child's exposure to SS and MF based on the county at birth. We repeated all analyses, recoding a child's exposure to SS and MF according the county where he or she attended third grade. According to the results presented in Table 5, the pattern of findings remained the same, and we can conclude that the findings are robust to different assumptions about exposure to SS and MF.

4. Discussion

The findings of this study show that allocations of state funds to counties for the Smart Start and More at Four (now NC Pre-k) programs are associated with positive outcomes on the population of targeted students at least through the end of middle school, with no evidence of fadeout or diminution of effect sizes across grades 3 through 8. For students in each middle school grade, early childhood exposure to higher levels of program funding is associated with higher average math and reading scores, a decreased likelihood of placement in special education, and a decreased probability of repeating a grade. These findings are consistent with findings from our prior studies of these two programs and show no evidence of fadeout during middle school. The pattern for Smart Start shows a stable positive association, and the pattern for More at Four shows an association of increasing magnitude across grade levels. These findings represent the longest-term evaluation ever

of statewide scaled-up early childhood programs, with the largest samples, using rigorous evaluation methods.

This evaluation was motivated in part by inconsistent past findings from other studies of early childhood programs, some of which reported that initially-favorable program effects faded out during elementary school years (Campbell & Ramey, 1994; Campbell et al., 2014; Heckman et al., 2010; Temple & Reynolds, 2007). Drawing from a consensus report that reviewed all existing studies (Dodge, 2017), we suggest that perhaps all reported findings from this and past studies should be treated as "true" and that different findings reflect different programs, populations, or school contexts. Early childhood programs are not identical.

Why might the two North Carolina programs demonstrate sustained positive associations when several other programs have effects that fade out? One explanation could be that these programs penetrated a large enough portion of the population to change the social ecology of children's educational experience such that the peer culture became a sustaining environment (Bailey et al., 2017; Dodge, 2018). MF provides funded slots for about a quarter of the population and spreads to more children in the population because the certified pre-kindergarten classrooms that enrolled funded students are also populated by nonfunded peers. SS improves the quality of early childhood education and care in the entire community by raising licensure standards and providing technical assistance and professional development to childcare providers; by design, it is intended to reach the entire population (Bryant et al., 2003). Both programs make every effort to ensure that young children are ready to learn at kindergarten entry; therefore, kindergarten teachers could accelerate teaching and students could focus on learning.

Other plausible explanations could be that the quality of the North Carolina programs is superior to that in other communities, or that the k-12 public education system in North Carolina works to sustain early growth in a not-yet specified way. In support of these explanations, the North Carolina Pre-K program was created by leaders of its k-3 school system and has been rated as highest quality according to published standards (Barnett, 2018). According to the annual report released in 2019 by National Institute for Early Education Research, North Carolina met 8 of 10 quality standards benchmarks such as a comprehensive early learning and development standard, curriculum supports, a maximum class size of 18, a low staff-child ratio of 1:9, a bachelor's degree and a birth-through-kindergarten license for lead teachers, a policy of screening and referral for students with health issues, and a continuous quality improvement system (Friedman-Krauss et al., 2019). In contrast, Florida, where pre-k positive program effects were not sustained, met only 2 of 10 quality standards benchmarks. Iowa preschool programs did not satisfy 4 quality standards such as degree requirement of lead teacher and assistant teacher, staff professional development, and a continuous quality improvement system. Tennessee's preschool system failed in degree requirement of assistant teacher, staff professional development, and a continuous quality improvement system.

The associations between funding for the More at Four pre-kindergarten program and middle school education outcomes varied significantly across subgroups within the

population. Positive associations emerged for economically advantaged and non-Black families but the positive associations were larger for children from more economically disadvantaged and Black families. These findings show that the targeting of economically disadvantaged children in the MF program had intended positive associations for this group while also having positive associations for other students.

Spillover associations on non-targeted students represent an added benefit of scaling up this program to reach the majority of the statewide population. About a quarter of the population was reached directly through funded slots, and more students were reached indirectly because they attended the same pre-kindergarten classrooms as the funded students and thus were exposed to the higher standards required for those classrooms to qualify for funding. Furthermore, given that many students had been enrolled in a certified pre-kindergarten experience, kindergarten teachers could begin the school year teaching to a higher level, thus benefitting additional students who had not been funded directly by MF funds or indirectly by being enrolled in certified classrooms.

Fig. 1 shows the benefit of SS is stable across grade levels, while the magnitude of MF effects increases across grades. We speculate that the difference comes from the more central role that MF plays in developing school readiness skills, as compared with SS. SS improves the quality of childcare services, which may help parents understand the value of schooling. In comparison to SS, MF focuses more on developing the foundational skills preschoolers should have when they enter kindergarten. Perhaps these skills beget the child's learning of more advanced skills during elementary school, which accelerates learning and increases the impact of MF, as Cunha and Heckman (2007, 2008) hypothesize.

4.1. Limitations and research implication

This study has several limitations worth noting. First, this study was an observational "natural experiment" design that allows for alternate explanations of findings. One possibility is that factors that led to variations in state funding across counties and years also led to variations in student outcomes, without the funding being causally responsible for the outcomes. We cannot generate a plausible pattern for such a case; that is, we cannot generate a narrative in which a factor that leads one county to obtain higher funding in one year but not other years would cause better outcomes for the children in that county-year cohort but not other cohorts, and would be the same factor replicated in other counties in other years. But because alternate conclusions remain possible, we remain cautious about causal conclusions.

Second, we are limited by a narrow range of available outcome variables. We did not examine non-cognitive skills of students. The formation of non-cognitive skills is as important as that of cognitive skills in child development. Thus, it is worth examining the early childhood program effects on non-cognitive skill formation such as social skills and ability of collaborative problem solving in future evaluation research.

Finally, a full cost-benefit analysis of SS and MF could inform policy. The reduction of grade repetition and special education services might lead to less spending, and improved

school performance contributes to higher educational achievements, which then lead to economic success in adulthood.

4.2. Policy implications

First, due to long-term favorable effects from Smart Start and More at Four on child education outcomes, we believe that North Carolina should continue supporting each program. The Smart Start Initiative and More at Four Program have continued for two decades in North Carolina. This evaluation is consistent with the conclusion that investment in early childcare and education can help a community achieve greater population-wide school success. To achieve this success, these programs should continue. In order to address the concerns from parents and their children aged birth to four, affordable and highquality early childcare supported by the Smart Start initiative needs to be accessible in all communities. High-quality pre-kindergarten for four-year-olds as currently implemented by the NC Pre-K program improves educational outcomes and should continue.

Second, the positive associations for each of these two programs supplement rather than substitute for each other. Funding one program only will not produce as strong an effect as funding both programs. One study limitation is that the evaluation of More at Four occurred in a context in which Smart Start was already present, so it is not clear what the effect of More at Four would be if Smart Start did not exist. The largest associations occur for those students who live in counties that receive relatively high allocations for each program. It would be a mistake to conclude that one program is sufficient to achieve maximum impact.

Third, policymakers must consider expanding early childhood program in order to serve additional eligible children. Fifty-two percent of children were eligible for NC Pre-K among 120,884 4-year-old children living in North Carolina in 2018. Unfortunately, 53 percent of these eligible children were not served due to insufficient budgets and resources to recruit and retain qualified teachers, expand facilities, and provide transportation (Barnett, 2018).

Fourth, North Carolina must review its policies on staff professional development for preschool teachers. According to the annual report released by National Institute for Early Education Research, North Carolina preschool teachers did not meet the requirement of paid time equivalent to their K-3 colleagues on planning, meeting, reporting, and professional development days (Friedman-Krauss et al., 2019).

Fifth, it is necessary to increase resources for evaluation and research. On one hand, we need to evaluate program implementation for quality control purposes because a program with sufficient funding does not guarantee effectiveness. On the other hand, outcome evaluations can help policymakers understand when and which sub-populations can be affected by the programs. Because our findings show that early childhood program effects did not fade out among middle school students, future research should examine adolescent outcomes such as high school dropout and graduation.

Finally, early childhood programs in other states can benefit from the successful experiences of SS and MF. Although funding resources, investment amount, school and community characteristics, and program structures differ across states, early childhood programs in

North Carolina represent models for other states. Our findings confirm positive effects of SS and MF program and add North Carolina to the list of the most successful state-funded early childhood programs in the U.S.

5. Conclusion

In summary, investments by the State of North Carolina in Smart Start and More at Four (NC Pre-k) are associated with long-term positive outcomes that last at least through the end of middle school with no fadeout over time. Compared to non-participants, children who were exposed to higher levels of funding for these early childhood services are better prepared to succeed academically in middle school. Many more students would benefit from participation in Smart Start and More at Four (NC Pre-K).

Acknowledgements

The authors thank the North Carolina Vital Records Office, the North Carolina Department of Public Instruction, and the North Carolina Education Research Data Center for providing data files. The authors acknowledge National Institutes of Health Grant R01 HD095930.

References

American Institutes for Research (2013). State-funded preschool: analyses of third-grade ITBS results of children enrolled in Iowa's 2007–08 Statewide voluntary preschool programs. Des Moines: Iowa, Department of Education.

Andrews RJ, Jargowsky P, & Kuhne K (2012). The effects of Texas's targeted prekindergarten program on academic performance (No. w18598).National Bureau of Economic Research.

Atteberry A, Bassok D, & Wong VC (2019). The effects of full-day pre-kindergarten: Experimental evidence of impacts on children's school readiness. Educational evaluation an policy analysis, 41, 537–562. 10.3102/0162373719872197

Bailey D, Duncan GJ, Odgers CL, & Yu W (2017). Persistence and fadeout in the impacts of child and adolescent interventions. Journal of Research in Educational Effectiveness, 10(1), 7–39.

Barnett W (2018). Barriers to expansion of NC Pre-K: Problems and potential solutions. http://nieer.org/wp-content/uploads/2019/01/NIEERNorthCarolina_2019.pdf.

Barnett W, Friedman-Krauss A, Weisenfeld G, Horowitz M, Kasmin R, & Squires J(2017). The state of preschool 2016.New Brunswick, NJ: National Institute for Early Education Research.

Barnett W, Jung K, Youn M, & Frede E (2013). Abbott preschool prooram longitudinal effects study: Fifth grade follow-up. New Brunswick, NJ: National Institute for Early Education Research.

Bryant D, Bernier K, Peisner-Feinberg E, & Maxwell K (2002). Smart start and child care in North Carolina: Effects on quality and changes over time. Chapel Hill: University of North Carolina, FPG Child Development Institute.

Bryant D, Maxwell K, Taylor K, Poe M, Peisner-Feinberg E, & Bernier K (2003). Smart Start and preschool child care quality in North Carolina: Change over time and relation to children's readiness. Chapel Hill: University of North Carolina.

Campbell F, Conti G, Heckman JJ, Moon SH, Pinto R, Pungello E, & Pan Y (2014). Early childhood investments substantially boost adult health. Science, 343 1478–1485. [PubMed: 24675955]

Campbell FA, & Ramey CT (1994). Effects of early intervention on intellectual and academic achievement: A follow-up study of children from low-income families. Child Development, 65, 684–698. 10.1111/j.1467-8624.1994.tb00777.x. [PubMed: 8013248]

Congressional Budget Office (2017). Options for reducing the deficit. https://www.cbo.gov/budget-options/2018/54778.

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

Cunha F, & Heckman JJ (2008). Formulating, identifying and estimating the technology of cognitive and noncognitive skill formation. Journal of human resources, 43(4), 738–782.

Dodge KA (Ed.). (2017). The current state of scientific knowledge on pre-kindergarten effects. Washington, DC: The Brookings Institution.

Dodge KA (2018). Toward population impact from early childhood psychological in terventions. American Psychologist, 73(9), 1117. [PubMed: 30525790]

Dodge KA, Bai Y, Ladd HF, & Muschkin CG (2017). Impact of North Carolina's early childhood programs and policies on educational outcomes in elementary chool Child Development 88(3), 996–1014. [PubMed: 27859011]

Early D, Li W, & Maxwell K (2017). Third-grade achievement for children who participated in Georgia's Pre-K. (Publication No. 2017–40). Bethesda, MD: Child Trends.

Follari L (2015). Foundations and best practices in early childhood education: History, theories, and approaches to learning. Pearson Higher Education AU.

Friedman-Krauss AH (2016). State(s) of Head Start: Funding, enrollment and quality are all over the map. http://nieer.org/2016/12/14/states-head-start-fundingenrollment-quality-map.

Friedman-Krauss AH, Barnett WS, Garver KA, Hodges KS, Weisenfeld GG, & DiCrecchio N (2019). The state of preschool 2018. State Preschool Yearbook. National Institute for Early Education Research.

Gamel-McCormick M, Amsden D, & Hartranft D (2005). Investing in better outcomes: Reaping continued dividends: The Delaware Early Childhood Longitudinal Study follow-up with fifth graders. Newark: University of Delaware, Center for Disabilities Studies.

Gormley WT Jr, 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.

Heckman JJ (2006). Skill formation and the economics of investing in disadvantaged children. Science, 312, 1900–1902. [PubMed: 16809525]

Heckman JJ, Moon SH, Pinto R, Savelyev PA, & Yvitz A (2010). The rate of return to the HighScope Perry Preschool Program. Journal of Public Economics, 94, 14–128.

Jung K, Barnett WS, Hustedt JT, & Francis J (2013). Longitudinal effects of the Arkansas Better Chance Program: Findings from first grade through fourth grade. New Brunswick, NJ: National Institute for Early Education Research, Rutgers University.

Ladd HF, Muschkin CG, & Dodge KA (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.

Leitner D (2013). Nevada Early Childhood Education (ECE) program: Building a foundation for school readiness and success in prek-12 and beyond: FY 2011–12 evaluation report. Carson City, Nevada: Office of Educational Opportunity.

Lipsey MW, Farran DC, & Durkin K (2018). Effects of the Tennessee pre-kindergarten program on children's achievement and behavior through third grade. Early Childhood Research Quarterly, 45, 155–176.

Maxwell K, Bryant D, & Miller-Johnson S (1999). A Six-County Study of the Effects of Smart Start Child Care on Kindergarten Entry Skills. Retrieved January 4, 2020, from https://fpg.unc.edu.

Miller LC, & Bassok D (2019). The effects of universal preschool on grade retention. Education Finance and Policy, 14(2), 149–177.

Muschkin CG, Ladd HF, & Dodge KA (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 CG, Ladd HF, Dodge KA, & Bai Y (2018). Gender differences in the impact of North Carolina's early care and education initiatives on student outcomes in elementary school. Educational Policy 0895904818773901.

North Carolina General Assembly (2019). DCDEE overview and base budget 2019 Retrieved Feburary 4, 2020, from https://www.ncleg.gov/documentsites/committees/JointAppropriationsHHS/2019%20Session/03-05-2019/DCDEE%200verview%20and%20Base%20Budget%202019.pdf.

North Carolina Partnership for Children (2019). Annual Report to the North Carolina General Assembly for Fiscal Year 2018–2019. Retrieved January 4, 2020, from http://www.smartstart.org/.

Peisner-Feinberg ES, & Schaaf JM (2006). Evaluation of the North Carolina More at Four Pre-kindergarten Program. Chapel Hill, NC: University of North Carolina.

Peisner-Feinberg ES, & Schaaf JM (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: FPG Child Development Institute.

Peisner-Feinberg ES, LaForett DR, Schaaf JM, Hildebrandt L, Sideris J, & Pan Y (2014). Chilaren's outcomes and program quality in the North Carolina Pre-Kindergarten Program: 2012–2013 statewide evaluation. Chapel Hill, NC: FPG Child Development Institute.

Peisner-Feinberg ES, Schaaf JM, Hildebrandt L, Pan Y, & Warnaar BL (2015). Children's kindergarten outcomes and program quality in the North Carolina PreKindergarten Program: 2013–2014 statewide evaluation. Chapel Hill, NC: FPG Child Development Institute.

Peisner-Feinberg ES, Garwood JD, & Mokrova IL (2016). Children's pre-k experiences and outcomes in the North Carolina pre-kindergarten program: 2014–2015 statewide evaluation. Chapel Hill, NC: FPG Child Development Institute.

Peisner-Feinberg ES, Mokrova LL, & Anderson T (2017). Effects of participation in the North Carolina pre-kindergarten program at the end of kindergarten: 2015–2016 statewide evaluation Chapel Hill NC: Child Development Institute

Ramey C, Ramey S, & Asmus G (2011). Longitudinal report: On-time 4th grade arrival of the LA 4 prekindergarten cohorts 1, 2 & 3. Lafayette: University of Louisiana at Lafayette, Cecil J. Picard Center for Child Development and Lifelong Learning.

Reynolds AJ, Ou S-R, Mondi CF, & Giovanelli A (2019). Reducing poverty and inequality through preschool-to-third-grade prevention services. American Psychologist, 74(6), 653–672. [PubMed: 31545639]

Schickedanz JA (1999). Much More than the ABCs: The Early Stages of Reading and Writing. NAEYC, 1509 16th Street, NW, Washington, DC 20036–1426 (order no. 204,$8).

Schweinhart LJ, Montie J, Xiang Z, Barnett WS, Belfield CR, & Nores M (2005), Lifetime effects: The High/Scope Perry preschool study through age 40 (Monographs of the High/Scope Educational Research Foundation). Ypsilanti, MI: High/Scope Educational Research Foundation.

Schweinhart LJ, Xiang Z, Daniel-Echols M, Browning K, & Wakabayashi T (2012). Michigan Great Start Readiness Program evaluation 2012: High school graduation and grade retention findings. Ypsilanti, MI: HighScope Educational Research Foundation

Shonkoff JP, & Phillips D (2000). From neurons to neighborhoods: The science of early child development. Washington, DC: National Academy Press.

Stern W (2017). Psychology of early childhood: Up to the sixth year of age. Routledge.

Temple JA, & Reynolds AJ (2007). Benefits and costs of investments in preschool education: Evidence from the Child-Parent Centers and related programs. Economics of Education Review, 26(1), 126–144. 10.1016/j.econedurev.2005.11.004.

U.S. Department of Health and Human Services, Administration for Children and Families (2010). Head Start Impact Study. Washington, DC: Final Report.

U.S. Government Accountability Office (2019). Child care and early education: Most States Offer Preschool Programs and Rely on Multiple Funding Sources. (Publication No. GAO-19–375).

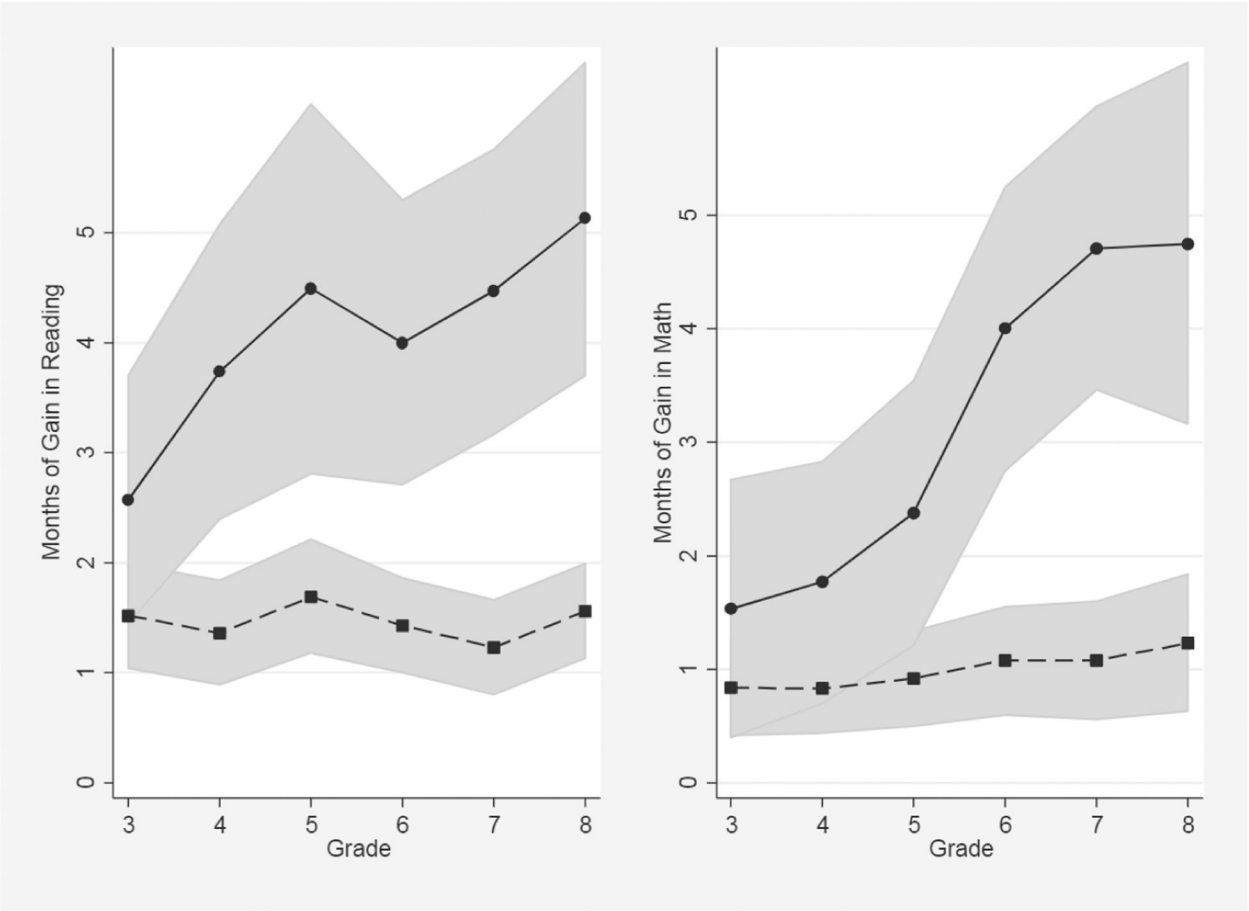

Fig. 1.

Additional months of learning for students exposed to the average funding levels of Smart Start and More at Four, by grade level. Note: Solid line represents More at Four. Dashed line represents Smart Start. Grey area represents 95% confidence interval. Gain scores are computed as the mean difference between students in counties at the average funded level and students in counties with no funding.

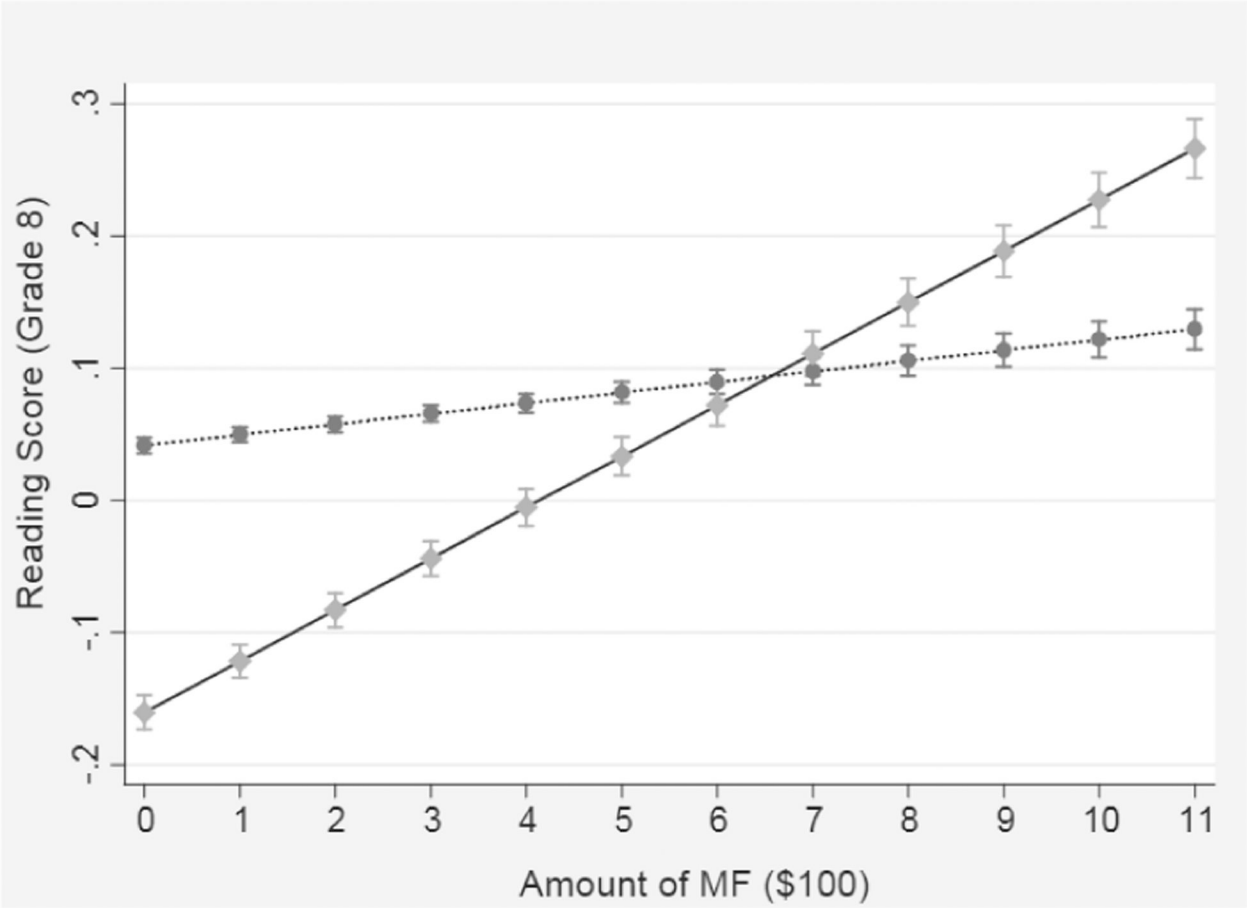

Fig. 2.

Interaction effects between More at Four program exposure and economically disadvantaged status of students on reading scores for eighth graders Note: Y axis is the mean standardized reading score. Solid line represents the group of economically disadvantaged students, defined as qualifying for free or reduced-price lunch. Dotted line represents the group of non-disadvantaged students. Scores are standardized with a population mean of 0 and standard deviation of 1.

Table 1

Mean academic performance by grade.

	Grade 6 (n = 907,738)			Grade 7 (n = 902,865)			Grade 8 (n = 896,349)		
Variables	**Mean**	**SD**	**N**	**Mean**	**SD**	**N**	**Mean**	**SD**	**N**
Academic outcomes									
Math standardized score	0.00	1.00	872,299	0.00	1.00	868,251	0.00	1.00	862,197
Reading standardized score	0.00	1.00	869,490	0.00	1.00	866,355	0.00	1.00	861,145
Grade retention	1.54%	–	907,738	1.45%	–	902,865	0.99%	–	896,349
Grade retention since G3	6.02%	–	907,738	7.04%	–	902,865	7.55%	–	896,349
Special education status	15.20%	–	907,738	14.70%	–	902,865	14.10%	–	896,349
Special education status since G3	19.75%	–	907,738	20.14%	–	902,865	20.37%	–	896,349

Table 2

Descriptive analysis.

	Grade 6 (n = 907,738)		
Variables	**Mean**	**SD**	**N**
Program			
Smart Start (non-zero, $00's)	11.36	8.62	770,887
More at Four (non-zero, $00's)	3.33	2.51	267,681
Smart Start ($00's)	9.67	8.92	905,130
More at Four ($00's)	0.99	2.04	905,130
Student characteristics			
Female	49.10%	–	907,733
Extremely low birth weight	0.46%	–	907,738
Very low birth weight	0.81%	–	907,738
Low birth weight	6.95%	–	907,738
Normal birth weight	81.80%	–	907,738
High birth weight	9.93%	–	907,738
Child white	60.50%	–	907,738
Child black	30.50%	–	907,738
Child native American	1.87%	–	907,738
Child Asian	0.97%	–	907,738
Child Hispanic	3.75%	–	907,738
Child mixed race	2.40%	–	907,738
Economic disadvantage	46.40%	–	906,675
Mother characteristics			
Mother's education (years)	12.52	2.41	906,582
Marital status	66.20%	–	907,693
Mother's age (years)	25.82	5.88	907,517
No dad information	14.70%	–	907,738
Mother immigrant	5.81%	–	907,642
First bom	44.00%	–	907,738
Mother white	63.20%	–	907,738
Mother black	30.40%	–	907,738
Mother native American	1.69%	–	907,738
Mother Asian	1.11%	–	907,738
Mother Hispanic	3.52%	–	907,738
Mother other race	0.06%	–	907,738
County-level demographic data, birth year			
Births to black mothers (share of births)	30.70%	16.90%	907,738
Births to Hispanic mothers (share of births)	3.51%	4.04%	907,738
Births to low education mothers (share of births)	23.70%	5.75%	907,738
Population on Food Stamps (share of population)	7.45%	3.77%	905,130
Population on Medicaid (share of population)	13.40%	5.66%	905,130

	Grade 6 (n = 907,738)		
Variables	**Mean**	**SD**	**N**
Number of births (log)	7.09	0.99	907,738
Total population (log)	11.70	0.99	905,130
Median family income (2009 $)	54,948	9922	905,130
School characteristics, test year			
Black students (share of students)	29.8	22.9	904,292
Other minority students (share of students)	12.7	11.7	904,292
Charter School	2.48%	15.60%	906,589
Per-pupil spending by source, test year			
Federal (2009 dollars)	844.5	468.2	898,944
State (2009 dollars)	4622	1248	898,944
Local (2009 dollars)	2170	1259	898,944
Same County	76.70%	–	907,738

Table 3

Regression models of smart start and more at four program impacts on educational outcomes.

	Grade 6	Grade 7	Grade 8
EOG math score			
Smart Start	0.0049*** (0.0011)	0.0049*** (0.0012)	0.0056*** (0.0014)
More at Four	0.0182*** (0.0029)	0.0214*** (0.0029)	0.0216*** (0.0037)
Observations	858,326	848,946	805,164
EOG reading score			
Smart Start	0.0065*** (0.0010)	0.0056*** (0.0010)	0.0071*** (0.0010)
More at Four	0.0182*** (0.0030)	0.0203*** (0.0030)	0.0233*** (0.0033)
Observations	855,538	847,607	804,051
Special education placement in each grade			
Smart Start	0.9916*** (0.0024)	0.9925*** (0.0021)	0.9873*** (0.0023)
More at Four	0.9607*** (0.0066)	0.9601*** (0.0061)	0.9502*** (0.0069)
Observations	893,102	884,039	833,592
Special education placement ever since Grade			
Smart Start	0.9911*** (0.0022)	0.9919*** (0.0020)	0.9876*** (0.0021)
More at Four	0.9663*** (0.0061)	0.9642*** (0.0057)	0.9594*** (0.0062)
Observations	893,102	884,039	833,592
Grade retention in each grade			
Smart Start	0.9959 (0.0074)	0.9974 (0.0069)	0.9998 (0.0072)
More at Four	0.9965 (0.0228)	0.9790 (0.0203)	0.9435*** (0.0199)
Observations	893,102	884,039	833,592
Grade retention ever since Grade 3			
Smart Start	0.9783*** (0.0078)	0.9765*** (0.0072)	0.9617*** (0.0060)
More at Four	0.9293*** (0.0178)	0.9263*** (0.0174)	0.8985*** (0.0186)
Observations	893,102	884,039	833,592

Note:

[a] All variables presented in Table 1 are controlled in the models.

[b] The unit of SS and MF investment are set to $100.

[c] Models use time and county fixed effects.

[d] Robust standard errors in parentheses.

[e] *** p<0.01, ** p<0.05, * p<0.1.

Table 4

Interactions between program effects and economic disadvantage of students.

	(1) Grade 6	(2) Grade 6	(3) Grade 7	(4) Grade 7	(5) Grade 8	(6) Grade 8
EOG reading score						
Smart Start	0.0065*** (0.0009)	0.0065*** (0.0011)	0.0056*** (0.001)	0.0058*** (0.0011)	0.0072*** (0.001)	0.0069*** (0.001)
More at Four	0.0182*** (0.003)	0.0135*** (0.0036)	0.0204*** (0.003)	0.0062* (0.0033)	0.0233*** (0.0033)	0.0032 (0.0031)
ED	−0.3309*** (0.0066)	−0.3377*** (0.0105)	−0.3170*** (0.0065)	−0.3324*** (0.0100)	−0.2969*** (0.0052)	−0.3336*** (0.0097)
SS x ED		0.0000 (0.0010)		−0.0005 (0.0010)		
MF x ED		0.0068** (0.0029)		0.0210*** (0.0030)		0.0319*** (0.0030)
EOG math score						
Smart Start	0.0049*** (0.0011)	0.0054*** (0.0014)	0.0049*** (0.0012)	0.0053*** (0.0014)	0.0058*** (0.0014)	0.0054*** (0.0016)
More at Four	0.0183*** (0.0030)	0.0145*** (0.0038)	0.0215*** (0.0029)	0.0032 (0.0032)	0.0216*** (0.0037)	−0.0012 (0.0033)
ED	−0.3387*** (0.0076)	−0.3354*** (0.0109)	−0.3253*** (0.0067)	−0.3429*** (0.0107)	−0.2973*** (0.0059)	−0.3394*** (0.0106)
SS x ED		−0.0009 (0.0011)		−0.0008 (0.0010)		0.0007 (0.0009)
MF x ED		0.0057* (0.0034)		0.0270*** (0.0031)		0.0363*** (0.0032)

Note:

[a] SS: Smart start; MF: More at Four; ED: economic disadvantage.

[b] The unit of SS and MF investment are set to $100.

[c] Reference group: students who were not identified as economic disadvantage.

[d] All variables presented in Table 1 are controlled in the models.

[e] Models use time and county fixed effects.

[f] Robust standard errors in parentheses.

g *** p<0.01, ** p<0.05, * p<0.1.

Table 5

Sensitivity analysis to examine program effects if a student's county of residence is different from the one where the student was born.

	(1) Full Sample	(2) Same County	(3) Different County	(4) Different County
EOG math score				
Grade 6				
Smart Start	0.0049*** (0.0011)	0.0061*** (0.0014)	0.0102*** (0.0027)	0.0316*** (0.0023)
More at Four	0.0182*** (0.0029)	0.0245*** (0.0040)	0.0045*** (0.0008)	0.0023*** (0.0009)
Observations	858,326	661,574	196,752	196,752
Grade 7				
Smart Start	0.0049*** (0.0012)	0.0061*** (0.0016)	0.0138*** (0.0028)	0.0316*** (0.0028)
More at Four	0.0214*** (0.0029)	0.0280*** (0.0040)	0.0042*** (0.0009)	0.0028*** (0.0009)
Observations	848,946	650,696	198,250	198,250
Grade 8				
Smart Start	0.0056*** (0.0014)	0.0068*** (0.0016)	0.0186*** (0.0029)	0.0286*** (0.0026)
More at Four	0.0216*** (0.0037)	0.0259*** (0.0047)	0.0047*** (0.0011)	0.0026*** (0.0008)
Observations	805,164	615,569	189,595	189,595
EOG Reading Score				
Grade 6				
Smart Start	0.0065*** (0.0010)	0.0078*** (0.0012)	0.0127*** (0.0027)	0.0303*** (0.0020)
More at Four	0.0182*** (0.0030)	0.0232*** (0.0041)	0.0052*** (0.0008)	0.0031*** (0.0007)
Observations	855,538	659,391	196,147	196,147
Grade 7				
Smart Start	0.0056*** (0.0010)	0.0069*** (0.0013)	0.0141*** (0.0030)	0.0312*** (0.0023)
More at Four	0.0203*** (0.0030)	0.0263*** (0.0040)	0.0044*** (0.0008)	0.0040*** (0.0007)
Observations	847,607	649,575	198,032	198,032
Grade 8				
Smart Start	0.0071*** (0.0010)	0.0082*** (0.0012)	0.0202*** (0.0032)	0.0254*** (0.0020)
More at Four	0.0233*** (0.0033)	0.0275*** (0.0040)	0.0056*** (0.0008)	0.0041*** (0.0006)
Observations	804,051	614,680	189,371	189,371

Note:

[a] All variables presented in Table 1 are controlled in the models.

[b] Models use time and county fixed effects.

[c] The unit of SS and MF investment are set to $100.

[d] Column (1) is the model output for full sample; Column (2) is the model for those if birth county was the same as school county; Column (3) and (4) are the models for those if birth county was different from school county: (3) uses $ from birth county; (4) uses $ from school county.

[e] Robust standard errors in parentheses.

[f] *** p<0.01, ** p<0.05, * p<0.1.