Article

Educational Policy
2020, Vol. 34(2) 377–407

DOI: 10.1177/0895904818773901
journals.sagepub.com/home/epx
SAGE

Gender Differences in the Impact of North Carolina's Early Care and Education Initiatives on Student Outcomes in Elementary School

Clara G. Muschkin[1], Helen F. Ladd[1], Kenneth A. Dodge[1], and Yu Bai[1]

Abstract
Based on growing evidence of the long-term benefits of enriched early childhood experiences, we evaluate the potential for addressing gender disparities in elementary school through early care and education programs. Specifically, we explore the community-wide effects of two statewide initiatives in North Carolina on gender differences in academic outcomes in Grades 3 to 5, using administrative student data and information on variation in program availability across counties and over time. We find that although investments in early care and education programs produce significant gains in math and reading skills on average for all children, boys experience larger program-related gains than girls. Moreover, the greatest gains among boys emerge for those from less advantaged families. In contrast, the large and statistically significant reductions in special education placements induced by these early childhood program do not differ consistently by gender.

Keywords
educational policy, gender, early childhood, preschool education, elementary education, achievement, finance

[1]Duke University, Durham, NC, USA

Corresponding Author:
Clara G. Muschkin, Sanford School of Public Policy, Duke University, Box 90545, Durham, NC 27713, USA.
Email: muschkin@duke.edu

Introduction

Despite prevailing norms of gender equity in education, boys and girls continue to perform differently in school. Although earlier evidence highlighted the stronger math skills of boys compared with girls, recent studies find clear advantages for girls, particularly in reading skills, that emerge early and persist across grade levels (Buchmann, DiPrete, & McDaniel, 2008; Fortin, Oreopoulos, & Phipps, 2015; Hyde, Lindberg, Linn, Ellis, & Williams, 2008; Loveless, 2015). Public perception of these gender differences has been framed as a "boy crisis" in public education, whereby school policies undermine male achievement across the ability spectrum and over time (Sommers, 2013). Policy makers and educators are particularly concerned about boys falling behind in reading, because these skills are considered pivotal for academic success at all grade levels, for attaining postsecondary education and for success in the labor market (Entwisle, Alexander, & Olson, 2007; Loveless, 2015). These concerns led to the federal *My Brother's Keeper* initiative, "to improve the educational futures of young African American and Hispanic boys, beginning in preschool and extending through high school graduation" (Rich, 2014; Thomas & Stevenson, 2009).

Gender disparities in academic performance, particularly as they persist across grade levels, may reflect differences in early learning and growth opportunities that result in lesser school readiness among boys (DiPrete & Jennings, 2012; Entwisle et al., 2007). Research has shown that outcomes in the early elementary grades are positively influenced by a child's access to enriched environments at early ages, through publicly funded preschool and early care programs, particularly among disadvantaged children (Duncan & Magnuson, 2013; Gormley & Gayer, 2005; Magnuson & Shager, 2010; Weiland & Yoshikawa, 2013).

Evidence is mixed, however, regarding the extent to which early education programs may mitigate gender differences at school entry and in the elementary grades. In their recent meta-analysis, Magnuson et al. (2016) find that across research studies, the effects of early childhood programs on cognitive academic outcomes are generally similar for girls and boys, with larger benefits for boys on school outcomes such as grade retention and special education placement. Although some individual program evaluation studies report greater benefits of early interventions for boys in the early grades (Deming, 2009; Hill, Gormley, & Adelstein, 2015; Joo, 2010; Ou & Reynolds, 2010), others report no gender differences (Anderson, 2008; Weiland & Yoshikawa, 2013).

Research Aims

We examine the differential benefits by gender of early care and education programs in North Carolina, which has for decades actively pursued early intervention policies to help children prepare for school and to reduce achievement gaps across grade levels. The state introduced and funded two statewide early childhood initiatives during our study period (1993-2010): *Smart Start* (SS) provided funding to improve child care services at the county level for all children between the ages 0 and 5 years; *More at Four* (MAF) provided funding for preschool slots for disadvantaged 4-year-olds. In prior research, we documented that investments in each of these programs generate significant educational gains for all groups of students in a community, and that these effects persist through the fifth grade (Dodge, Bai, Ladd, & Muschkin, 2016; Ladd, Muschkin, & Dodge, 2014; Muschkin, Ladd, & Dodge, 2015). In the present study, we examine the extent to which investments in early childhood programs can improve educational outcomes for boys and reduce gender disparities in the elementary grades. We do so by evaluating how, over time, state investments in North Carolina's two early childhood initiatives affect gender differences in math, reading, and special education placement in Grades 3 through 5. Based on recent research evidence that gender differences are exacerbated by socioeconomic disadvantage, we also ask whether the potential mitigating effects of early education on gender differences in academic outcomes may differ by family background (Autor, Figlio, Karbownik, Roth, & Wasserman, 2016; Garcia, Heckman, & Ziff, 2017; Magnuson et al., 2016).

Our research contributes to the literature on early gender differences in three ways. First, our focus on the potential effects of programs that intervene prior to school entry extends the extant literature on gender disparities downward to early childhood. In our analyses of elementary school outcomes, we ask specifically whether access to early enrichment, prior to kindergarten entry, is particularly beneficial for boys' preparedness for success throughout elementary school. Second, we use a novel approach for evaluating long-term impacts of early childhood programs on student outcomes. By taking advantage of the variation across North Carolina counties and years in the timing of the introduction and funding levels of two programs, we are able to examine the combined effects of the programs on both the participating and the nonparticipating students, and, thus, contribute to a broader understanding of the patterns and malleability of gender disparities in the early grades. This strategy permits us to avoid some selection issues that challenge the validity of many studies evaluating the benefits of participating in early childhood interventions (Dodge et al., 2016). A third contribution of this study is to the small but growing body of research evidence on gender-differentiated benefits of early education. A more complete understanding of these differences may help educators and policy makers harness the potential of early intervention to address a growing gender gap in lifetime educational achievement (Owens, 2016).

Background

In framing the relationship of gender differences to early education, we begin with a brief overview of trends in academic performance by gender in the United States and North Carolina, and a review of the research that examines the relationship of these gender differences to the development of early skills and behaviors across family and educational environments. We then review empirical studies that inform our understanding of mechanisms through which early childhood education programs may reduce subsequent gender disparities. Finally, we describe the two North Carolina early childhood programs and the characteristics that suggest that they might effectively help boys to overcome early disadvantage.

Trends in Gender Differences

Test score trends on national standardized assessments (the National Assessment of Education Progress [NAEP]) document academic performance patterns over time, as reported in Figures 1 and 2, for the United States and North Carolina. Trends in fourth-grade NAEP scores over time indicate an increasing gender similarity in math and a consistent female advantage in reading, both in the United States and in North Carolina. In North Carolina, boys' and girls' math achievement converged early, so that by 2011, there was no mean gender difference. In contrast, boys' disadvantage in reading scores has remained steady over time, with less narrowing of the gender gap over time in North Carolina as compared with the U.S. trend. In North Carolina, in 2011, girls averaged 8 points higher than boys in reading NAEP scores. This difference translates to an average fourth-grade gender gap of more than half of a grade level (Hyde et al., 2008).

Sources of Gender Disparities in Education

In efforts to explain gender differences in educational outcomes, researchers have joined forces across disciplines to study a multiplicity of influences, including biological factors, social and economic factors, and educational

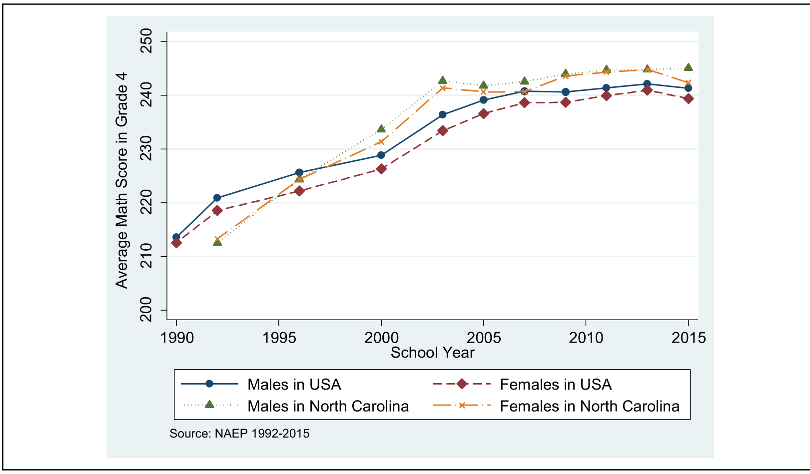

Figure 1. Trends in math scores, fourth-grade NAEP.
Note. NAEP = National Assessment of Education Progress.

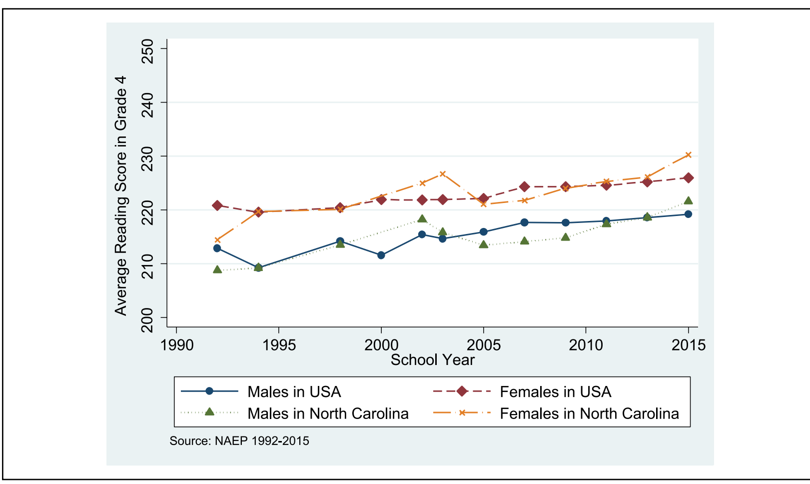

Figure 2. Trends in reading scores, fourth-grade NAEP.
Note. NAEP = National Assessment of Education Progress.

policy and practices. Although biological sex differences may be related to educational outcomes at school entry through mechanisms such as differences in motor skill and speech development, sex differences in cognition are small and inconsistent (Spelke, 2005). Studies exploring physiological impacts of prenatal environmental and maternal stress on cognitive development and educational outcomes do not find consistent variation by gender (Autor et al., 2016). One physiological factor is clearly gender differentiated: the consistent overrepresentation of boys among children identified with chronic disabilities and genetic abnormalities that are likely to be present at birth. This gender disparity is often highlighted in the analyses of risk of special education placement in school (DiPrete & Jennings, 2012; Rutter et al., 2004).

The focus of the research literature on gender and education policy has recently shifted away from specific educational practices that disproportionately disadvantage either boys or girls, in favor of approaches that "map a complex terrain of current gender inequalities" (Buchmann et al., 2008). Recent studies highlight the importance of examining gender disparities in conjunction with family background influences. Studies that report academic advantages for girls note that these are greatest among disadvantaged students—those receiving meal subsidies, English language learners, and minority students (Entwisle et al., 2007; Fryer & Levitt, 2010; Loveless, 2015).

In addition to conceptualizing gender differences in the context of other sources of educational disadvantage, gender studies in education have shifted their focus away from the later grades to examining disparities in academic performance in early and middle childhood. Recent studies provide compelling evidence that early school experiences are key to understanding gender as one of many different factors associated with long-term educational inequities (Buchmann et al., 2008; DiPrete & Jennings, 2012; Entwisle et al., 2007; Husain & Millimet, 2009). For both boys and girls, early interventions provided in high-quality settings may provide important educational foundations, compensating for a potential lack of early intellectual stimulation in less advantaged homes (Dodge & Haskins, 2015; Hart & Risley, 1995). Analyses of gender gaps in reading and math outcomes suggest differences in the development of early cognitive skills and the multidimensional set of social and behavioral skills that determine school readiness. Levels of these skills at early ages have been shown to affect a variety of long-term educational outcomes, and research finds that they are more developed among girls than boys at school entry (DiPrete & Jennings, 2012; Duncan et al., 2007; Jones, Greenberg, & Crowley, 2015; Ou & Reynolds, 2010).

Another potential explanation for gender differences in educational achievement comes from an emerging literature suggesting that boys are more susceptible to some environmental influences than are girls (Broekhuizen, van Aken, Dubas, Mulder, & Leseman, 2015; Susman, 2006; Votruba-Drzal, Coley, Maldonado-Carren, Li-Grining, & Chase-Lansdale, 2010). Boys are conceived at a higher rate than girls but suffer greater morbidity across the life span, due to greater susceptibility to environmental hazards. This perspective asserts that adverse influences such as low income and racial discrimination will have greater impact on boys than girls. Support for this conclusion comes from empirical findings of greater gender differences in educational outcomes among lower income children (Autor et al., 2016). These authors find that the negative effects of early family disadvantage are more profound for boys as compared with girls, and that this difference persists from kindergarten entry through age 18. This differential susceptibility hypothesis applies equally well to positive environmental influences, suggesting that early childhood experiences of enrichment through high-quality early education programs may have greater positive impact on boys than girls and could mitigate gender differences in educational achievement.

How Can Early Childhood Programs Reduce Gender Disparities?

By providing opportunities for cognitive stimulation and social and emotional growth, high-quality early childhood programs can mitigate risks associated with disadvantaged early environments. Evaluation studies across many types of programs find that the educational and lifetime benefits of high-quality early childhood education are most pronounced among children from economically disadvantaged homes (Barnett, 2011; Garcia et al., 2017; Gormley & Gayer, 2005; Jones et al., 2015; Phillips et al., 2017; Dodge et al., 2016; Ladd et al., 2014; Muschkin et al., 2015). Few studies of early interventions have analyzed the heterogeneity of effects by both family disadvantage and gender (Entwisle et al., 2007; Wood, Kaplan, & McLoyd, 2007). An exception is a recent analysis of experimental data from the Carolina Abecedarian Project. Garcia et al. (2017) find that gender differences in long-term program impacts vary with the quality of early environments. Boys attending low-quality child care programs experienced detrimental lifetime outcomes, whereas girls were less vulnerable to this condition.

Early education programs also may improve educational outcomes for boys through well-balanced curricula that can mitigate the selective gender role orientations that may lead boys toward less enriching types of early childhood activities. Early gender role socialization may determine later preferences and practices that carry over to school readiness in language and numeracy skills (Orr, 2011). Gender role reinforcement may lead to a preference among boys for games that are physically interactive, competitive, and less directly supervised by adults, whereas girls are encouraged to engage in activities that develop fine motor skills, verbal interactions, and higher levels of creativity (Early et al., 2010; Orr, 2011). Superior skills in the “feminine” areas are linked to the large and increasing gender gaps observed in reading scores across grade levels (Entwisle et al., 2007; Fortin, Oreopoulos, & Phipps, 2015; Loveless, 2015). With regard to math skills, research suggests that gender role orientation encourages boys more than girls to foster early interest in activities involving numbers, patterns, and shapes. Children's math attitudes begin to form in preschool, with girls displaying lower math self-concepts than boys (Gunderson, Ramirez, Levine, & Beilock, 2012; Jacobs, Lanza, Osgood, Eccles, & Wigfield, 2002; Lubienski, Robinson, Crane, & Colleen, 2013; Robinson & Lubienski, 2011). However, trends in NAEP scores belie the conventional wisdom that boys experience an advantage in math skills when they enter school. Instead, we observe a convergence of math skill levels during the early elementary grades in the United States and to a greater extent in North Carolina. This trend may reflect the effects of high-quality early childhood programs with well-balanced curricula on reducing gendered behaviors and attitudes that inhibit math achievement among girls. These gendered behaviors and attitudes are more prevalent among children of lower income families (Buchmann et al., 2008; Entwisle et al., 2007). Thus, we might expect more and persistent gender differences in math and reading skills among children from less advantaged families.

Early education programs can help to mitigate boys' disadvantage in the development of specific social skills and behaviors that allow girls to enter school more ready than boys to succeed. Expectations for elementary-grade classroom behavior are consistent with behaviors that gender socialization emphasizes for girls, such as following directions, rule learning, help seeking, cooperation, and prosocial peer interactions (Duckworth & Seligman, 2006; Orr, 2011). These are the behaviors that can be cultivated and strengthened in high-quality early education environments.

With regard to special education placements, we consider the role of early education within a developmental perspective, whereby disabilities are a function of a child's endowment and experiences early in life (Caspi & Moffit, 2006). Consistent with that perspective, in a prior study, we found that early childhood programs did not influence placements for severe conditions or chronic disabilities, but did significantly reduce the likelihood of placements for learning disability, mild mental handicaps, developmental delays, and language disorders. We note that the latter categories are the most malleable to gains in the school readiness skills associated with positive early education experiences (Muschkin et al., 2015). Boys have higher rates of the severe and chronic exceptionality conditions, suggesting that the early interventions may have less pronounced effects of reducing the gender gap in special education placements, as compared with the potential impact on math and reading scores.

Early Education Initiatives in North Carolina

North Carolina has actively pursued early childhood policies, providing support for designing and implementing two statewide initiatives during our study period (1993-2010). *Smart Start* provided funding to improve child care services at the county level for all children between the ages 0 and 5 years, and *More at Four* provided funding for preschool slots for disadvantaged 4-year-olds. Both SS and MAF began as demonstration projects in a small number of counties and were later expanded to serve all counties.

Smart Start

The Smart Start initiative began in 1993 as a state-funded program for comprehensive early childhood development, administered by the nonprofit North Carolina Partnership for Children. The partnership provides funds to county nonprofit organizations, to be used for a variety of early childhood services such as day care vouchers, health and family support services, child care services, professional development, and promoting collaborations among local agencies. One of the goals of SS is to increase the proportion of high-quality child care centers, where early enrichment and intervention to prevent developmental delays is most likely to occur. Organizations have substantial discretion over how they use SS funds, provided they use the funds to promote the development of children between the ages 0 and 5 years.

Across counties, average funding per 0- to 5-year-old peaked in 2000 at about US$400 per child and fell to about US$220 per child in 2009 (the last year of current evaluation). A child living in a county with Smart Start funding for all five of his or her early childhood years would have access to Smart Start funding equal to about 5 times these per child amounts. Both the penetration and funding levels varied considerably over time and across counties.

The state does not set income level eligibility requirements for SS, but it does require that at least 30% be spent on child care services. The flexibility and breadth of the Smart Start initiative leads us to consider its broader impact on children: Rather than considering SS a narrowly defined program that has an impact only on its direct participants, we evaluate its system-level effort to improve the quality of early childhood environments for the entire community. The potential for addressing multiple sources of early disadvantage suggests that the SS initiative can help to reduce gender disparities during the preschool years. Through investments in developmental screening, fostering high-quality curricula in early education programs and monitoring child care quality, Smart Start funds can contribute to improvements in cognitive and social skills among all children, with even greater benefits for boys.

More at Four

The More at Four program began in 2001 with the goal of improving school readiness among at-risk children in the year prior to entering kindergarten. Four-year-olds were eligible for MAF if the family had annual income at or below 75% of the state median income, or if the child was limited in English proficiency, disabled, chronically ill, or had a developmental need. In addition, MAF served a critical role in ensuring screening for all special needs and providing access to needed health and social services through referrals to other agencies within the county.

Funding for MAF was provided as classroom-based slots for eligible children, usually in existing preschool programs. Across the study period, the program provided an average of US$4,200 per slot per year, which often was supplemented with other funding sources such as Title I or Individuals With Disabilities Education Act (IDEA) Part B for special education interventions provided within the classroom. Over the period of our study, approximately half of the MAF slots were in public school settings, a third in for-profit community child care centers, and the remainder in nonprofit child care centers and Head Start programs. To qualify for these state-funded slots, classrooms had to meet specific quality standards for staff qualifications, class size, teacher–child ratios, and state child care licensing requirements.

Because many children not funded by More at Four dollars were enrolled in the same centers as those who received funds, they also benefited from the spillover of high-quality standards required for MAF classrooms. Program evaluators found that these classroom quality standards were consistently maintained throughout the process of scaling-up MAF over this period, enrolling approximately 25% of all 4-year-olds in 2010 (N.C. Office of State Budget and Management, 2014; Peisner-Feinberg & The More at Four Evaluation Team, 2008). MAF has contributed to statewide improvements in pre-K quality: According to National Institute for Early Education Research (NIEER) national evaluations of preschool programs, quality standards in North Carolina increased from a score of 7 in 2003 to 10 (the highest possible score) in 2010 (Barnett et al., 2010). In recognition of these and other potential spillovers, we use the same approach as for Smart Start and define the "treatment" as the availability and level of More at Four funding per age-appropriate child in the county. All counties in North Carolina provided funding for More at Four by 2004. Throughout the study period, there was significant variation in the distribution of funding and availability of classroom slots across counties (Ladd et al., 2014).

MAF's specific focus on school readiness and on children who are developmentally at high risk of problems in school suggests that this program might target the preacademic and social skills that originate prior to school entry and that are less developed among boys. Furthermore, as the program targets at-risk children, it can help boys from poor families to overcome disadvantages resulting from less enriched early home environments. The structure of MAF funding and its impact on classroom quality means that the program may generate positive effects for all children that extend beyond the benefits for individual participants. Thus, consistent with our earlier studies, we examine the effects of both SS and MAF on improving academic outcomes for all age-appropriate children in a community, with a focus on additional benefits for boys.

Data

We explore the community-wide effects of SS and MAF on gender differences in academic outcomes among children enrolled in third grade between 1995 and 2010, in fourth grade between 1996 and 2011, and in fifth grade between 1997 and 2012. For each of the analyses, we use administrative student data and information on variation, across North Carolina counties and over time, in the availability and penetration of these programs.

Information about students and the schools they attend in the third grade is drawn from administrative data files from the North Carolina public schools. To identify the county in which a child was born and eligible to receive early education services, we link the education data to individual birth records for all children born in the state between 1988 and 2000 and later enrolled in third, fourth, and fifth grades between 1995 and 2012. The sample includes only the children whose records we were able to link across domains—that is, students who were born in North Carolina and were enrolled in a public elementary school in a North Carolina county. Of the 1,347,562 births during the study period, 1,004,571 were matched as attending a public or charter school at some time in Grades 3, 4, or 5 (74.55% of children born in North Carolina during this period). We use the birth record data to identify the control variables we include in our models: gender, child's race and ethnicity, birth weight, marital status of the mother, education level of the mother, whether the father is identified, and the race, ethnicity, and immigrant status of the mother.

The outcome variables in our models are drawn from the educational administrative data files that are submitted by each school district to the North Carolina Department of Public Instruction (NCDPI) and made available by the Duke University North Carolina Education Research Data Center. These files contain student-level information that has been linked longitudinally so that educational outcomes can be tracked over time. Math and reading scores on the end of grade tests in third, fourth, and fifth grades are provided by the North Carolina DPI Division of Accountability, and information on special education placements is submitted by the Exceptional Children Division. A child was coded as having a special education placement in each grade if any category of exceptionality was noted (except for academically gifted).

A third data set contains administrative records of funding levels by county by year, for Smart Start and More at Four. All funding levels are calibrated to 2009 U.S. dollars. The average SS investment for a child was about US$1,100 (US$220 per year across 5 years from birth through age 4). The average MAF investment was about US$330 per 4-year-old (countywide average, even though only about one fourth of students were enrolled); however, this average cost includes the years before MAF was initiated in 2001. In the years when MAF was fully funded, the MAF investment per 4-year-old in the population was about US$1,100, similar to the SS investment per child aged 0 to 5 years. Information on school characteristics and district expenditures is drawn from the National Center for Education Statistics (NCES) Common Core of Data. Publicly available information is used to identify county-level demographic characteristics that vary over time.

Analysis Strategy

Our first step is to identify the direction and magnitude of program impact on math and reading scores and on special education placements in third through fifth grades, by gender. As in our previous studies, we test the effects *for all children* in the affected counties of investments in these two statewide initiatives in North Carolina. Our approach is to estimate average program effects as a function of state financial allocations to the county in which children were born and attend school, in the year(s) when each child was age eligible for that funding. We specify ordinary least squares (OLS) models for the test score outcomes and logistic regression models to predict special education placement. The basic models include program investments and gender as the main variables of interest in each model, and the full models include program-by-gender interaction terms, to test for heterogeneity of program effects. To evaluate the extent to which gender differences in program effects may vary by family disadvantage, we estimate the full models, including gender/program interactions, in separate analyses for two groups defined by the mother's education level.

The full models for outcomes in Grades 3, 4, and 5 take the following forms:

$$O_{icbtg} = \beta_1 SS_{icb}^* + \beta_2 MAF_{icb^*} + \beta_3 X_{ib} + \beta_4 Y_{it} + \beta_5 D_{icb} + \beta_6 F * MAF_{icb} + \beta_7 F * SS_{icb} + \alpha_c + \gamma_b + \varepsilon_{icbt}, \quad (1)$$

$$\text{Logit}\left[P\left(E_{icbtg} = 1\right)\right] = \beta_1 SS_{icb}^* + \beta_2 MAF_{icb}^* + \beta_3 X_{ib} + \beta_4 Y_{it} + \beta_5 D_{icb} + \beta_6 F * MAF_{icb} + \beta_7 F * SS_{icb} + \alpha_c + \gamma_b + \varepsilon_{icbt}, \quad (2)$$

where O_{icbtg} and Logit($P[E_{icbtg} = 1]$) are Grade g outcome (g = 3, 4, and 5; math and reading scores, special education status, respectively) in year t for the ith student born in county c in year b. SS_{icb}* is a Smart Start variable for the ith child, as defined further below. MAF_{icb}* is a More at Four variable for the ith child, as defined further below. $F*MAF_{icb}$ is the interaction term between gender and the MAF variable for the ith child. $F*SS_{icb}$ is the interaction term between gender and the SS variable for the ith child. $\mathbf{X}_{ib}$ is a vector of characteristics of the ith child and his or her parents children at the time of birth including the child's gender, child's birth weight, and the education level of the child's mother; $\mathbf{Y}_{it}$ are characteristics of the ith child observed in year t, such as the race of the child; D_{icb} are county demographic characteristics corresponding to the child's birth year; α_c and γ_b are county and year fixed effects, where year refers to the year of birth.

Of primary interest are the coefficients of the two program variables, as well as their interactions with the gender variable. In the basic model, we define the Smart Start variable as $SS_{icb}^* = \sum PSS_{icb(\text{age} = 0, \ldots 4)}$, where b refers to the period of eligibility for SS funding. PSS refers to the penetration of Smart Start defined as the inflation-adjusted dollars per child aged 0 to 5 years in the relevant county in each of the years when the child was below 5 years. As the sum of the penetration rates over the years when the child was below 5, this measure accounts not only for different spending per year but also for the possibility that funding may not have been available in the relevant county during all the child's preschool years.

The More at Four penetration measure MAF_{icb}* is defined as $PMF_{icb(\text{age} = 4)}$. In this case, the penetration rate refers to the inflation-adjusted dollars per 4-year-old for More at Four in the relevant county for the year in which the child was 4 years old. This variable is normalized by the total number of 4-year-olds in a county, not by the smaller number of 4-year-olds who meet the "at-risk" eligibility criterion. Consequently, one explanation for the variation across counties in this measure is the variation in the proportion of at-risk 4-year-olds to all 4-year-olds. Much of the variation of this type would be absorbed, however, by the fixed effects terms that are included in all models: county fixed effects (α_c) and year fixed effects (γ_b) that adjust for time-invariant county differences and secular trends in outcomes. Thus, program coefficients can be interpreted as effects of SS and MAF funding within counties across time and across counties within years.

Results: Descriptive Analysis of Educational Outcomes by Gender

Table 1 provides descriptive statistics for all outcome, predictor, and control variables. Table 2 summarizes, for each grade, unadjusted mean gender differences in student outcomes by mother's education level (low level equals less than 12 completed years, and comprises 23% of the study sample). Throughout the analyses, we use mother's education level as a proxy measure of family socioeconomic status (SES) because we do not have access to student-level information on household income or parental occupation. Educational attainment of mothers is available from each child's birth record. We consider this to be an appropriate measure in the absence of more complete information, based on findings that parental education is a better predictor of education outcomes than is family income (Davis-Kean, 2005).

In Table 2, grade-level mean differences in outcomes favor girls in all but Grade 3 math scores, and are statistically significant. The slight male advantage in math in Grade 3 is no longer significant in Grades 4 and 5, whereas the female advantage in reading scores, as well as in special education placements, remains constant across grades. Figures 3 and 4 illustrate gender differences by subgroups of students defined by lower versus higher mother's education level. Comparing gender differences across subgroups and grades, the female advantage in reading scores and special education status is largest among students with less educated mothers. However, it is clear in both sets of charts that the most pronounced differences in math, reading, and special education placements are between rather than within these SES subgroups. Among fourth graders, boys in the less advantaged group scored 0.61 standard deviations lower than boys in the more advantaged group. The difference between less advantaged boys and more advantaged girls is 0.77 standard deviations. To put this difference in context, the typical annual gain in test scores from the beginning to the end of Grade 4 is about 0.5 standard deviations in reading. If we assume a 10-month school year, a difference of 0.1 standard deviations represents about 2 months of instruction in reading (and 1.5 months in math). The mean differences in reading scores, thus, indicate that higher SES girls are 15 months, or 1.5 grade levels, ahead of lower SES boys. A consistent pattern emerges for special education placements: Less advantaged boys are 9.6% more likely than other boys to be placed in

Table 1. Descriptive Statistics.

	Grade 3 (*n* = 906,194)		Grade 4 (*n* = 910,744)		Grade 5 (*n* = 910,527)	
Variables	*M*	*SD*	*M*	*SD*	*M*	*SD*
Academic outcomes						
Math standardized test score	0.0	1.0	0.0	1.0	0.0	1.0
Reading standardized test score	0.0	1.0	0.0	1.0	0.0	1.0
Special education placement (%)	15.0	35.7	15.7	36.4	15.6	36.3
Program funding						
SS (US$00s)—only years with nonzero allocations	11.5	8.6	11.5	8.6	11.4	8.6
MAF (US$00s)—only years with nonzero allocations	3.3	2.5	3.3	2.5	3.3	2.5
SS (US$00s)—all years	9.9	8.9	9.8	8.9	9.7	8.9
MAF (US$00s)—all years	1.0	2.1	1.0	2.1	1.0	2.0
Student characteristics (%)						
Female	49.1	50.0	49.1	50.0	49.1	50.0
Extremely low birth weight	0.5	6.8	0.5	6.8	0.5	6.7
Very low birth weight	0.80	8.9	0.8	8.9	0.8	8.9
Low birth weight	6.9	25.4	7.0	25.4	6.9	25.4
Normal birth weight	81.9	38.5	81.9	38.6	81.9	38.5
High birth weight	9.9	29.9	9.9	29.9	9.9	29.9
Child White	60.9	48.8	60.8	48.8	60.6	48.8
Child Black	30.2	45.9	30.2	45.9	30.3	46.0
Child native American	1.8	13.4	1.8	13.4	1.9	13.5
Child Asian	1.0	10.0	1.0	10.0	1.0	9.9
Child Hispanic	3.6	18.7	3.6	18.7	3.7	18.8
Child mixed race	2.5	15.5	2.5	15.7	2.6	15.8
Mother characteristics						
Mother's education (years)	12.5	2.4	12.5	2.4	12.5	2.4
Mother's low education (<12 years)	23.3	42.3	23.4	42.3	23.3	42.3
Marital status (% married)	66.4	47.2	66.4	47.2	66.3	47.2
Mother's age (years)	25.85	5.89	25.84	5.89	25.83	5.88
No dad information (%)	14.5	35.2	14.5	35.3	14.6	35.3
Mother immigrant (%)	6.0	23.7	5.9	23.6	5.9	23.5
First born (%)	44.0	49.6	44.0	49.6	44.0	49.6
Mother White (%)	63.4	48.2	63.4	48.2	63.3	48.2
Mother Black (%)	30.1	45.9	30.2	45.9	30.2	45.9
Mother native American (%)	1.7	12.8	1.7	12.8	1.7	12.8
Mother Asian (%)	1.2	10.7	1.1	10.6	1.1	10.6
Mother Hispanic (%)	3.6	18.7	3.6	18.6	3.5	18.5
Mother other race (%)	0.1	2.3	0.1	2.4	0.1	2.3
County-level demographic data by birth year						
Births to Black mothers (% share of births)	30.1	16.6	30.4	16.7	30.4	16.7
Births to Hispanic mothers (% share of births)	3.6	4.1	3.6	4.1	3.5	4.0
Births to low education mothers (% share of births)	23.4	5.8	23.6	5.8	23.5	5.8
Population on food stamps (% share of population)	7.5	3.8	7.5	3.8	7.5	3.8
Population on Medicaid (% share of population)	13.4	5.7	13.4	5.7	13.4	5.7
Number of births (log)	7.075	1	7.092	0.996	7.085	0.995
Total population (log)	11.71	0.988	11.71	0.987	11.71	0.986
Median family income (2009 US$)	57,699	10,451	57,669	10,436	57,639	10,417

(continued)

Table 1. (continued)

Variables	Grade 3 (*n* = 906,194) *M*	*SD*	Grade 4 (*n* = 910,744) *M*	*SD*	Grade 5 (*n* = 910,527) *M*	*SD*
School characteristics, test year						
Black students (% share of students)	30.1	24.2	29.8	24.1	29.3	24.1
Other minority students (share of students)	11.9	12.7	12.7	13.0	13.7	13.4
Per-pupil spending by source, test year						
Federal (2009 U.S. dollars)	638	338	722	392	802	423
State (2009 U.S. dollars)	4,172	1,050	4,074	1,178	3,961	1,300
Local (2009 U.S. dollars)	2,307	1,422	2,623	1,614	2,939	1,783
Attended school in the county of birth (%)	78.8	40.9	78.0	41.4	77.3	41.9

Note. SS = Smart Start; MAF = More at Four.

special education in fourth grade, and 19% more likely than more advantaged girls. The pattern of male disadvantage, particularly among lower income boys, persists across Grades 3 through 5.

Results: Estimated Effects of Early Education Programs by Gender

We estimate the potential for early education programs to mitigate gender differences within and between subgroups, through regression models in which the primary coefficients of interest are for indicators of gender, mother's education, and the two program variables, measured as units of US$100 investments in SS and MAF. These are per-child expenditures, over the years that the child was eligible for services within the county of the student's birth and school enrollment. The program coefficients in the models are adjusted for the student characteristics, mother attributes, and county and school characteristics summarized in Table 1.

Table 3 summarizes the results of regression models for math and reading scores in each grade (Equation 1), for both the basic models that include the program and gender variables and the full models that add program-by-gender interaction terms. Coefficients for selected variables are reported in Table 3, and the full sets of coefficients are reported in Online Appendix Table 1.

The main effects coefficients for MAF and SS in Table 3 indicate that program investments contributed positively and significantly to math and reading scores for boys and girls, across grades. We use the fourth-grade basic model (the third and ninth columns of Table 3) to translate the coefficients to effect sizes at a meaningful funding level. Using the coefficients in columns 3 and 9, we estimate that the average investment in SS (US$1,100 or US$220 per 0- to 5-year-old child per each of 5 years of eligibility of

Table 2. Mean Gender Difference by Mother's Education Level.

	Total			Mother low education			Mother high education		
Academic outcomes	Male *M*	Female *M*	*p* for *t* test	Male *M*	Female *M*	*p* for *t* test	Male *M*	Female *M*	*p* for *t* test
Grade 3									
Math standardized score	0.01	−0.01	***	−0.45	−0.46	**	0.14	0.13	***
Reading standardized score	−0.08	0.08	***	−0.56	−0.38	***	0.06	0.22	***
Special education status	19.90%	9.90%	***	26.40%	13.80%	***	17.90%	8.71%	***
Grade 4									
Math standardized score	0.00	0.00	>.05	−0.46	−0.45	**	0.13	0.13	>.05
Reading standardized score	−0.08	0.08	***	−0.55	−0.39	***	0.06	0.22	***
Special education status	20.70%	10.60%	***	28.10%	15.10%	***	18.50%	9.15%	***
Grade 5									
Math standardized score	0.00	0.00	>.05	−0.47	−0.45	***	0.13	0.14	>.05
Reading standardized score	−0.07	0.07	***	−0.55	−0.40	***	0.07	0.21	***
Special education status	20.60%	10.50%	***	28.40%	15.30%	***	18.20%	9.01%	***

**$p < .05$. ***$p < .01$.

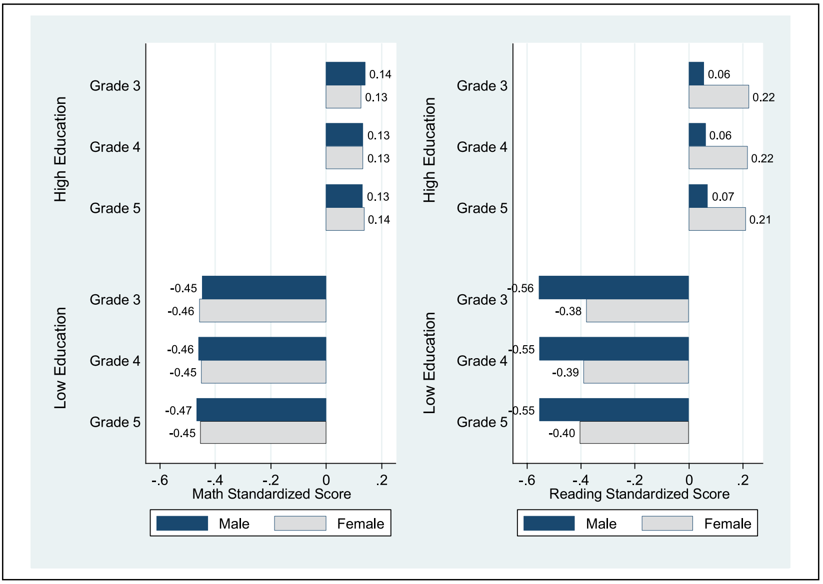

Figure 3. Math and reading outcomes by gender and mother's education.

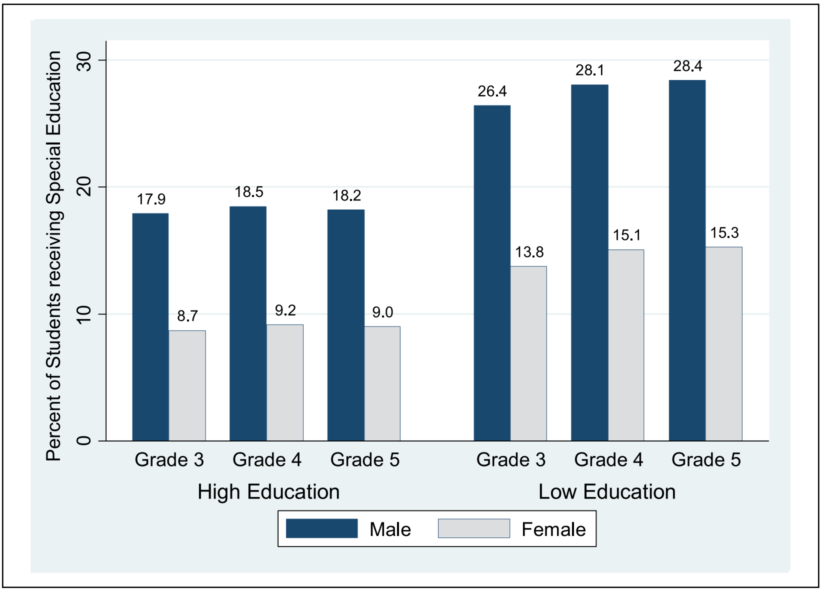

Figure 4. Special education outcomes by grade and mother's education.

Table 3. OLS Regression Models for Math and Reading Scores, With and Without Gender Interaction Terms.

	Math scores						Reading scores					
	Grade 3 basic	Grade 3 gender interaction	Grade 4 basic	Grade 4 gender interaction	Grade 5 basic	Grade 5 gender interaction	Grade 3 basic	Grade 3 gender interaction	Grade 4 basic	Grade 4 gender interaction	Grade 5 basic	Grade 5 gender interaction
SS (US$00s)	0.0046***	0.0055***	0.0047***	0.0053***	0.0053***	0.0062***	0.0065***	0.0068***	0.0058***	0.0062***	0.0074***	0.0078***
	(0.0013)	(0.0013)	(0.0012)	(0.0012)	(0.0013)	(0.0013)	(0.0011)	(0.0011)	(0.0011)	(0.0011)	(0.0011)	(0.0012)
MAF (US$00s)	0.0100***	0.0115***	0.0150***	0.0168***	0.0188***	0.0202***	0.0126***	0.0164***	0.0179***	0.0225***	0.0208***	0.0223***
	(0.0035)	(0.0035)	(0.0034)	(0.0035)	(0.0035)	(0.0037)	(0.0027)	(0.0028)	(0.0030)	(0.0033)	(0.0037)	(0.0039)
Female	0.0003	0.0211***	0.0159***	0.0313***	0.0133***	0.0333***	0.1801***	0.1938***	0.1691***	0.1861***	0.1467***	0.1583***
	(0.0035)	(0.0050)	(0.0035)	(0.0044)	(0.0033)	(0.0045)	(0.0034)	(0.0041)	(0.0032)	(0.0039)	(0.0036)	(0.0041)
Mother's low education (<12 years)	−0.2462***	−0.2462***	−0.2463***	−0.2463***	−0.2395***	−0.2395***	−0.2461***	−0.2461***	−0.2404***	−0.2404***	−0.2335***	−0.2335***
	(0.0053)	(0.0053)	(0.0054)	(0.0054)	(0.0051)	(0.0051)	(0.0053)	(0.0052)	(0.0052)	(0.0052)	(0.0048)	(0.0048)
Female × SS		−0.0018***		−0.0012***		−0.0018***		−0.0006*		−0.0008**		−0.0009**
		(0.0003)		(0.0003)		(0.0003)		(0.0003)		(0.0004)		(0.0004)
Female × MAF		−0.0029**		−0.0034***		−0.0026**		−0.0075***		−0.0089***		−0.0028**
		(0.0012)		(0.0012)		(0.0012)		(0.0011)		(0.0012)		(0.0012)
Constant	−4.1294*	−4.1426*	−5.3386**	−5.3433**	−5.7288**	−5.7377**	−4.0898***	−4.1016***	−3.3946**	−3.4000**	−4.2456**	−4.2511**
	(2.3168)	(2.3138)	(2.1206)	(2.1177)	(2.4850)	(2.4818)	(1.4793)	(1.4775)	(1.6355)	(1.6332)	(1.8135)	(1.8123)
Observations	868,119	868,119	865,873	865,873	834,117	834,117	864,348	864,348	861,272	861,272	830,096	830,096
R^2	.236	.237	.238	.238	.236	.236	.230	.230	.238	.238	.238	.238

Note. Robust standard errors in parentheses. Each column reports selected coefficients from full models that also include student characteristics measured at birth and at each grade time-varying county attributes, school attributes, and county and year fixed effects (county and year of birth). Online Appendix Table 1 provides full information on this model. OLS = ordinary least squares; SS = Smart Start; MAF = More at Four.

*$p < .1$. **$p < .05$. ***$p < .01$.

funding) generated an average gain in fourth-grade math scores of 0.052 standard deviations (the basic model SS coefficient, 0.0047 × 11) and in Grade 4 reading scores of 0.064 (model coefficient 0.0058 × 11). The effect size for this average level of SS investment is equivalent to one instructional month gain in both reading and math scores. The equivalent calculation for the effect of MAF on fourth-grade outcomes reveals that the program led to a gain of 0.2 standard deviations or four instructional months in reading, and a gain of 0.165 standard deviations in math or 2.5 months of instruction. In this additive model, access to this level of investment in both programs leads, on average, to a combined program effect (SS + MAF) of five instructional months in reading and 3.5 months in math in the fourth grade.

To look at gender effects, we include both a gender indicator variable and program-by-gender interaction terms in the regression models predictive of math and reading scores in the third, fourth, and fifth grades. The first set of analyses includes the basic model with gender and mother's education level as main predictors of interest, and a second model that brings in the program-by-gender interaction terms. Unlike the unadjusted mean differences displayed in Table 2, in the basic model predictive of math scores (column 1, Table 3), female gender is not a significant predictor in Grade 3 but is associated with significantly higher math scores in Grades 4 and 5 (columns 3 and 5). The female effect on reading is positive and significant across all grades, as depicted in the second panel of Table 3. For all outcomes and across grades, the higher level of mother's education has a positive and significant impact on math and reading scores.

The program-by-gender interaction terms in the regression models for each grade in Table 3 indicate whether the effects of the program variables on math and reading scores vary significantly by gender. With the interaction term in the model, the main effect of each program variable refers to its effect on the math and reading scores of the reference category (boys). The negative and significant coefficients of the interaction terms across grades indicate that, on average, funding for each program in both reading and math had a bigger impact on boys than on girls.

We also explored the extent to which variation in program effects by gender differs by SES subgroups, defined in our study by mother's education level. Table 4 summarizes the full models that incorporate the gender-by-program interaction term, estimated separately for two groups: students whose mothers have less than a high school degree and students whose mothers have at least a high school degree. Across models, the program effects are positive and significant on math and reading scores, for all subgroups and grades. Comparing the pattern of coefficients of the students whose mothers have less education with those of students whose mothers have more education, we note that girls in both groups are at an advantage in math scores in

Table 4. OLS Regression Models for Math and Reading Scores, With Gender Interactions, by Subgroup of Mother's Education Level.

	Math scores						Reading scores					
	Grade 3 mom high education	Grade 3 mom low education	Grade 4 mom high education	Grade 4 mom low education	Grade 5 mom high education	Grade 5 mom low education	Grade 3 mom high education	Grade 3 mom low education	Grade 4 mom high education	Grade 4 mom low education	Grade 5 mom high education	Grade 5 mom low education
SS (US$00s)	0.0053***	0.0067***	0.0052***	0.0059***	0.0056***	0.0083***	0.0066***	0.0074***	0.0062***	0.0067***	0.0074***	0.0093***
	(0.0014)	(0.0015)	(0.0012)	(0.0014)	(0.0014)	(0.0015)	(0.0010)	(0.0014)	(0.0010)	(0.0015)	(0.0011)	(0.0017)
MAF (US$00s)	0.0104***	0.0154***	0.0148***	0.0242***	0.0189***	0.0255***	0.0158***	0.0195***	0.0212***	0.0277***	0.0208***	0.0283***
	(0.0037)	(0.0040)	(0.0036)	(0.0041)	(0.0037)	(0.0047)	(0.0029)	(0.0037)	(0.0035)	(0.0042)	(0.0039)	(0.0052)
Female	0.0188***	0.0283***	0.0300***	0.0352***	0.0294***	0.0466***	0.1907***	0.2038***	0.1833***	0.1948***	0.1555***	0.1680***
	(0.0056)	(0.0066)	(0.0051)	(0.0061)	(0.0050)	(0.0070)	(0.0045)	(0.0063)	(0.0046)	(0.0070)	(0.0042)	(0.0065)
Female × SS	−0.0015***	−0.0027***	−0.0012***	−0.0013**	−0.0015***	−0.0028***	−0.0004	−0.0011**	−0.0006	−0.0016**	−0.0007*	−0.0018***
	(0.0004)	(0.0005)	(0.0003)	(0.0006)	(0.0004)	(0.0005)	(0.0004)	(0.0005)	(0.0004)	(0.0006)	(0.0004)	(0.0006)
Female × MAF	−0.002**	−0.0032*	−0.0030**	−0.0046**	−0.0025*	−0.0031	−0.0070***	−0.0091***	−0.0092***	−0.0082***	−0.0030**	−0.0020
	(0.0013)	(0.0019)	(0.0012)	(0.0022)	(0.0013)	(0.0023)	(0.0012)	(0.0018)	(0.0013)	(0.0021)	(0.0013)	(0.0022)
Constant	−4.1294*	−4.1426*	−5.3386**	−5.3433**	−5.7288**	−5.7377**	−4.0898***	−4.1016***	−3.3946**	−3.4000**	−4.2456**	−4.2511**
	(2.3168)	(2.3138)	(2.1206)	(2.1177)	(2.4850)	(2.4818)	(1.4793)	(1.4775)	(1.6355)	(1.6332)	(1.8135)	(1.8123)
Observations	868,119	868,119	865,873	865,873	834,117	834,117	864,348	864,348	861,272	861,272	830,096	830,096
R^2	.236	.237	.238	.238	.236	.236	.230	.230	.238	.238	.238	.238

Note. Robust standard errors in parentheses. Each column reports selected coefficients from full models that also include student characteristics measured at birth and at each grade, time-varying county attributes, school attributes, and county and year fixed effects (county and year of birth). Online Appendix Table 2 provides full information on this model. OLS = ordinary least squares; SS = Smart Start; MAF = More at Four.
$^*p < .1$. $^{**}p < .05$. $^{***}p < .01$.

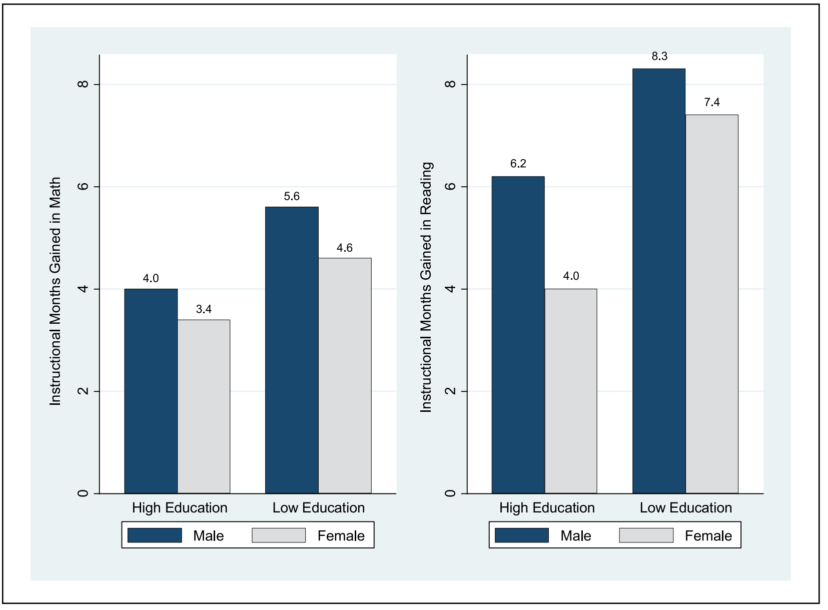

Figure 5. Combined SS and MAF program effects (instructional months) on Grade 5 reading and math scores, by gender and mother's education level.
Note. SS = Smart Start; MAF = More at Four.

fourth and fifth grades, and in reading in all grades. For each subgroup and grade, the negative sign of the program-by-gender interaction term indicates that boys experienced more gains in math and reading associated with SS and MAF investments—with a few exceptions, these coefficients are statistically significant across grades.

Program effect sizes for each grade level are summarized in Figure 5 and detailed in Online Appendix Table 5, in terms of the average number of instructional months gained in reading and math, from a US$1,100 investment in SS and in MAF. At each grade level and for both subgroups, the positive SS and MAF effects on math and reading scores are greater for boys than for girls. For example, in fourth grade, boys gained an additional 2 months in reading on average, as compared with girls; boys in the lower education group experienced the largest reading gain of 6.1 months. The combined effects of both programs for Grade 5 are depicted in Figure 5, which illustrates variation in program impact, in terms of months gained in reading and math, by gender and by mother's education level. This set of results summarizes important dimensions of the community-wide effects of early childhood programs: Benefits across all student groups in terms of reading and math

Table 5. Logistic Regression Models for Special Education Placement in Grades 3 to 5, With and Without Gender Interaction Terms.

	Grade 3 basic	Grade 3 gender interaction	Grade 4 basic	Grade 4 gender interaction	Grade 5 basic	Grade 5 gender interaction
SS (US$00s)	0.9914***	0.9911***	0.9928***	0.9922***	0.9943**	0.9934***
	(0.0025)	(0.0024)	(0.0025)	(0.0024)	(0.0024)	(0.0023)
MAF (US$00s)	0.9723***	0.9721***	0.9532***	0.9517***	0.9447***	0.9418***
	(0.0066)	(0.0066)	(0.0063)	(0.0063)	(0.0064)	(0.0067)
Female	0.4288***	0.4251***	0.4363***	0.4272***	0.4346***	0.4210***
	(0.0032)	(0.0039)	(0.0027)	(0.0035)	(0.0027)	(0.0034)
Female × SS		1.0008		1.0017*		1.0025***
		(0.0009)		(0.0009)		(0.0009)
Female × MAF		1.0006		1.0044		1.0080**
Constant	687.7448*	689.4159*	79.6290	80.2397	7.5169	7.6404
	(2,359.3738)	(2,365.0006)	(252.9634)	(254.9048)	(28.4525)	(28.9454)
Observations	894,386	894,386	898,311	898,311	896,423	896,423

Note. Entries are odds ratios. Robust standard errors for the estimated log of the odds coefficients are in parentheses. Each column reports selected coefficients from full models that also include student characteristics measured at birth and at each grade time-varying county attributes, school attributes, and county and year fixed effects (county and year of birth). Online Appendix Table 3 provides full information on this model. SS = Smart Start; MAF = More at Four.
**p* < .1. ***p* < .05. ****p* < .01.

gains are substantial and significant; in each subgroup, boys benefit from the programs to a greater extent than girls; and students from less advantaged families gain the most from access to these programs with boys benefiting more so than girls.

The logistic regression models corresponding to Equation 2 predict special education placement in third, fourth, and fifth grade and are summarized in Table 5, which reports odds ratios for selected variables (the full set of odds ratios is reported in Online Appendix Table 3). In the basic models in Table 5, allocations for MAF and SS significantly reduce the likelihood that a student will have an exceptionality classification across all grades, as indicated by odds ratios <1 for both program variables. Also across grades, the gender gap in special education placement is large and significant: Boys were, on average, more than twice as likely as girls to have a special education placement (odds ratios < 0.50 for the variable "female"). In contrast to the full OLS models (Table 3), the gender-by-program interaction terms are mostly nonsignificant in the logistic regression models in Table 5, indicating that program benefits do not vary consistently by gender. Similarly, most of these interaction terms are not statistically significant across the models summarized in Table 6, estimated separately for students with higher and lower educated mothers.

Table 6. Logistic Regression Models for Special Education Placement in Grades 3 to 5, With Gender Interactions, by Subgroup of Mother's Education Level.

	Grade 3 mom high education gender interaction	Grade 3 mom low education gender interaction	Grade 4 mom high education gender interaction	Grade 4 mom low education gender interaction	Grade 5 mom high education gender interaction	Grade 5 mom low education gender interaction
SS (US$00s)	0.9905***	0.9920***	0.9919***	0.9926***	0.9935**	0.9932**
	(0.0026)	(0.0029)	(0.0027)	(0.0028)	(0.0026)	(0.0027)
MAF (US$00s)	0.9708***	0.9736***	0.9503***	0.9535***	0.9374***	0.9500***
	(0.0065)	(0.0086)	(0.0062)	(0.0092)	(0.0059)	(0.0105)
Female	0.4237***	0.4258***	0.4243***	0.4307***	0.4187***	0.4236***
	(0.0047)	(0.0080)	(0.0044)	(0.0074)	(0.0048)	(0.0065)
Female × SS	1.0007	1.0014	1.0017	1.0021	1.0025**	1.0028*
	(0.0011)	(0.0016)	(0.0011)	(0.0014)	(0.0011)	(0.0015)
Female × MAF	1.0023	0.9967	1.0053	1.0035	1.0081*	1.0084
Constant	545.5004*	1,946.3610*	50.0553	530.6159	20.7783	0.9676
	(1,992.8989)	(7,569.3884)	(163.8526)	(2,086.5530)	(81.9430)	(4.3614)
Observations	685,000	209,386	687,642	210,669	685,984	210,439

Note. Robust standard errors for the estimated log of the odds coefficients are in parentheses. Each column reports selected coefficients from full models that also include student characteristics measured at birth and at each grade time-varying county attributes, school attributes, and county and year fixed effects (county and year of birth). Online Appendix Table 3 provides full information on this model. SS = Smart Start; MAF = More at Four.
**p* < .1. ***p* < .05. ****p* < .01.

Once again, we use fourth-grade models (from Table 5) to illustrate the magnitude of these effects, by calculating the effect size for each program at the average funding level of US$1,100, which is 11 times the program investment variable unit of US$100. In the basic model for Grade 4, the significant SS odds ratio of 0.9928 indicates that this average investment reduced the probability of special education placement by 8% (based on the odds ratio for this investment level and calculated as exp(ln[0.9928] × 11). The odds of being placed in special education for a student born in a county with an MAF investment of US$1,100 per child was 0.59 in Grade 4 (exp(ln[0.9532] × 11), indicating that at this level of funding, MAF reduced the probability of special education placement by 41% in Grade 4. The reduction in special education placements associated with each program also was very similar for students in the two subgroups defined by mother's education level (Table 6). Taken together, this pattern of results indicates that the early education programs have an important beneficial effect on reducing special education placements in the elementary grades, with effects that are similar for boys and girls and within and across subgroups of students from less and more advantaged families.

Results: Sensitivity Analyses

In each of our evaluations of early interventions in North Carolina, we have addressed concerns about the potential bias that could emerge from the process of matching birth to student records. We have chosen to include in our sample all children for whom we could match birth to student records, regardless of whether the county of birth was the same as the child's place of residence when enrolled in Grade 3. Each student's SS and MAF funding allocations are set to correspond to the county of birth, although in fact some students may have attended school in a different county. To determine whether the program effects are sensitive to this assumption, we have estimated the models predictive of Grade 3 math and reading scores on two alternative samples: children who lived in the same county at birth and in Grade 3 and children who moved to a different country between birth and Grade 3. Although the effect sizes for the two programs were slightly larger in the same-county sample than in our original analyses, and slightly smaller in the different-county sample, all coefficients remained significant across assumptions. These sensitivity analyses are described in detail in Dodge et al. (2016).

In the current study, comparing our matched sample with the population of elementary school students, we find a slight overrepresentation in our sample of children whose mother is non-White, a single parent, and immigrant. Our strategy for addressing this issue is to include all these child characteristics covariates in each model, and find that these characteristics are not related to a child's access to program funding.

Conclusion and Policy Implications

This study addresses the extent to which investments in high-quality early education programs may mitigate achievement differences by gender in middle childhood. To better understand these patterns and how to target them through policy, we also ask to what extent are gender differences related to family background, and can early childhood program investments reduce gender disparities among the least advantaged students? We first approach these questions by exploring, for the elementary grades, the patterns of variation in academic outcomes by gender. Focusing on reading, math, and special education placement among 13 cohorts of third, fourth, and fifth graders in North Carolina (Figures 3 and 4), we note that boys are at a substantial disadvantage in reading and special education placements, and are at a slight disadvantage in math. Across all outcomes, the most pronounced disadvantages are present among boys with less educated mothers. These findings contribute to the research evidence that gender differences are present early in the education process, and that these differences persist as children progress through school. Furthermore, the findings underscore the need to understand the multiple sources of risk that may contribute to academic underperformance among boys.

Our findings confirm that investments in high-quality early childhood education programs have significant and beneficial effects for both boys and girls in each of the elementary grades, in terms of reading and math skills, as well as in reducing special education placements. For example, in the fourth grade, average program effect sizes, in terms of instructional months in reading gained through program funding at the average levels, are on the order of 1 month for SS and 4 months for MAF. Compared with findings from other evaluation studies, the average effect sizes for SS are similar to those of community programs such as in Tulsa and New Jersey, whereas the average effects of MAF are larger than those of most pre-K programs. These effect sizes reflect our research design; whereas most evaluation studies measure program impact only for children who directly participated in programs, in our study, we estimate effects for all children in the community who were age eligible for the programs. The MAF effect sizes reported in our analyses are substantial, particularly considering that they are averaged across all children, which suggests that the estimates of program benefits involve large spillovers to nonparticipant children. These comprehensive benefits are consistent with the goals of both programs. The SS initiative aims to improve early childhood environments across many domains, which would potentially help all children to improve school readiness skills. The MAF program, though targeted to at-risk children, used a funding strategy that led to improvements in pre-K classrooms that could benefit children who did not directly participate in the program. Moreover, it is likely that spillover benefits contribute to improved classroom environments in the elementary grades, when children are in classes with peers who are better prepared for school.

We find that program investments have the potential to reduce gender disparities in early achievement levels: At each grade level and across SES groups, boys receive the most pronounced program benefits. Combining the estimated effects for Smart Start and More at Four, we find that fourth-grade boys experienced an additional reading gain of two instructional months as compared with girls. Boys from less advantaged families had the largest gain of all subgroups, gaining six instructional months (more than half of a grade level in reading). Our study does not provide information on the mechanisms through which these early education programs promote student success, but we suggest that the comprehensive nature of the SS program and the potential spillovers of the MAF program have led to enhancement of early environments, both at school and prior to kindergarten, in ways that are particularly helpful for boys.

Although investments in both SS and MAF substantially reduce special education placements across the elementary grades, the program effects did not vary consistently by gender. To a certain extent, this finding is not surprising. Enriched early environments can help children through more effective screening for developmental issues, which would benefit boys and girls in equal measure. Improved access to developmental screening is an explicit goal of the Smart Start initiative. Many special education placements are based on diagnoses of chronic conditions and genetic issues that are more common among boys, however, and early education programs would not alter these conditions. Nonetheless, it is possible that for some diagnostic categories and for the least severe conditions, early childhood programs may reduce disproportionate special education placement among boys by fostering school readiness skills that allow a healthy transition to kindergarten. A more detailed analysis of exceptionality diagnosis and severity is beyond the scope of this study but would be useful for better understanding this important educational gender disparity.

This study builds on our earlier findings of widespread population-level benefits linked to funding early childhood programs, as measured by individual student outcomes across the elementary grades. By estimating specific and comprehensive benefits, we can contribute information that is directly useful for policy makers regarding funding for early education. We conclude that policies that enhance early childhood environments for children are an effective means of addressing the disadvantages that many students face when they are entering school and in the early grades. This study provides evidence of another important set of returns to investing in early childhood programs: providing a more gender equitable education for all children.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

References

Anderson, M. (2008). Multiple inference and gender differences in the effects of early intervention: A reevaluation of the Abecedarian, Perry Preschool, and Early Training projects. *Journal of the American Statistical Association*, *103*, 1481-1495. doi:10.1198/016214508000000841

Autor, D., Figlio, D., Karbownik, K., Roth, J., & Wasserman, M. (2016). *Family disadvantage and the gender gap in behavioral and educational outcomes* (NBER

Working Paper No. 22267). Cambridge, MA: National Bureau of Economic Research.

Barnett, W. S. (2011). Effectiveness of early educational intervention. *Science*, *333*, 975-978.

Barnett, W. S., Epstein, D. J., Carolan, M. E., Fitzgerald, J., Ackerman, D. J., & Friedman, A. H. (2010). *The state of preschool 2010: State preschool yearbook*. New Brunswick, NJ: National Institute for Early Education Research.

Broekhuizen, M. L., van Aken, M. G., Dubas, J. S., Mulder, H., & Leseman, P. M. (2015). Individual differences in effects of child care quality: The role of child affective self-regulation and gender. *Infant Behavior & Development*, *40*, 216-230. doi:10.1016/j.infbeh.2015.06.009

Buchmann, C., DiPrete, T. A., & McDaniel, A. (2008). Gender inequalities in education. *Annual Review of Sociology*, *34*, 319-337.

Caspi, A., & Moffitt, T.E. (2006). Gene-environment interactions in psychiatry: Joining forces with neuroscience. *Nature Reviews Neuroscience*, *7*, 583-590.

Davis-Kean, P. E. (2005). The influence of parent education and family income on child achievement: The indirect role of parental expectations and the home environment. *Journal of Family Psychology*, *19*, 294-304.

Deming, D. (2009). Early childhood intervention and life-cycle skill development: Evidence from Head Start. *American Economic Journal: Applied Economics*, 111-134. doi: 10.1257/app.1.3.111

DiPrete, T. A., & Jennings, J. L. (2012). Social and behavioral skills and the gender gap in early educational achievement. *Social Science Research*, *41*, 1-15.

Dodge, K. A., & Haskins, R. (2015). Children and government. In R. M. Lerner, M. H. Bornstein, & T. Leventhal (Eds.), *Handbook of child psychology and developmental science, volume four, ecological settings and processes* (7th ed., pp. 654-703). New York, NY: John Wiley.

Dodge, K. A., Bai, Y., Ladd, H. F., & Muschkin, C. G. (2016). Impact of North Carolina's early childhood programs and policies on educational outcomes in elementary school. *Child Development*, *88*, 996-1014.

Duckworth, A. L., & Seligman, M. E. P. (2006). Self-discipline gives girls the edge: Gender differences in self-discipline, grades, and achievement test scores. *Journal of Educational Psychology*, *98*, 198-208.

Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., . . . Japel, C. (2007). School readiness and later achievement. *Developmental Psychology*, *43*, 1428-1446.

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

Early, D. M., Iruka, I. U., Ritchie, S., Barbarin, O. A., Winn, D. C., Crawford, G. M., . . . Pianta, R. C. (2010). How do Pre-Kindergarteners spend their time? Gender, ethnicity, and income as predictors of experiences in Pre-Kindergarten classrooms. *Early Childhood Research Quarterly*, *25*, 177-193.

Entwisle, D. R., Alexander, K. L., & Olson, L. S. (2007). Early schooling: The handicap of being poor and male. *Sociology of Education*, *80*, 114-138.

Fortin, N. M., Oreopoulos, P., & Phipps, S. (2015). Leaving boys behind: Gender disparities in high academic achievement. *Journal of Human Resources*, *50*, 549-579.

Fryer, R. G., & Levitt, S. D. (2010). An empirical analysis of the gender gap in mathematics. *American Economic Journal: Applied Economics*, *2*, 201-240.

Garcia, J. L., Heckman, J. J., & Ziff, A. L. (2017). *Gender differences in the benefits of an influential early childhood program* (NBER Working Paper No. 23412). Cambridge, MA: National Bureau of Economic Research.

Gormley, W. T., Jr., & Gayer, T. (2005). Promoting school readiness in Oklahoma: An evaluation of Tulsa's pre-k program. *Journal of Human Resources*, *40*, 533-558.

Gunderson, E. A., Ramirez, G., Levine, S. C., & Beilock, S. L. (2012). The role of parents and teachers in the development of gender-related math attitudes. *Sex Roles*, *66*, 153-166. doi:10.1007/s11199-011-9996-2

Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children*. Baltimore, MD: Brookes Publishing.

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

Husain, M., & Millimet, D. L. (2009). The mythical "boy" crisis? *Economics of Education Review*, *28*, 38-48.

Hyde, J. S., Lindberg, S. M., Linn, M. C., Ellis, A. B., & Williams, C. C. (2008). Gender similarities characterize math performances. *Science*, *321*, 494-495.

Jacobs, J. E., Lanza, S., Osgood, D. W., Eccles, J. S., & Wigfield, A. (2002). Changes in children's self-competence and values: Gender and domain differences across grades one through twelve. *Child Development*, *73*, 509-527.

Jones, D. E., Greenberg, M., & Crowley, M. (2015). Early social-emotional functioning and public health: The relationship between kindergarten social competence and future wellness. *American Journal of Public Health*, *105*, 2283-2290. doi:10.2105/AJPH.2015.302630

Joo, M. (2010). Long-term effects of Head Start on academic and school outcomes of children in persistent poverty: Girls vs. boys. *Children and Youth Services Review*, *32*, 807-814. doi:10.1016/j.childyouth.2010.01.018Kirp

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*, 162-187.

Loveless, T. (2015). Girls, boys, and reading. *Brown Center Report on American Education*, *3*, 3-25.

Lubienski, S. T., Robinson, J. P., Crane, C. C., & Colleen, M. G. (2013). Girls' and boys' mathematics achievement, affect, and experiences: Findings from ECLS-K. *Journal for Research in Mathematics Education*, *44*, 634-645.

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.

Magnuson, K. A., & Shager, H. (2010). Early education: Progress and promise for low-income children. *Children and Youth Services Review*, *32*, 1186-1198. doi:10.1016/j.childyouth.2010.03.006

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*, 478-500

North Carolina Office of State Budget and Management. (2014). *County/state population projections*. Retrieved from https://www.osbm.nc.gov/facts-figures/demographics

Orr, A. J. (2011). Gendered capital: Childhood socialization and the "boy crisis" in education. *Sex Roles*, *65*, 271-284.

Ou, S., & Reynolds, A. J. (2010). Mechanisms of effects of an early intervention program on educational attainment: A gender subgroup analysis. *Children and Youth Services Review*, *32*, 1064-1076. doi:10.1016/j.childyouth.2009.10.020

Owens, J. (2016). Early childhood behavior problems and the gender gap in educational attainment in the United States. *Sociology of Education*, *89*, 236-258. doi:10.1177/0038040716650926

Peisner-Feinberg, E. S., & The More at Four Evaluation Team. (2008). *Evaluation of the North Carolina more at four Pre-Kindergarten Program, year 6*. Chapel Hill: The University of North Carolina, Frank Porter Graham Child Development Institute.

Phillips, D. A., Lispsey, M. A., Dodge, K. A., Haskins, R., Bassok, D., Burchinal, M. R., . . . Weiland, C. (2017). *Puzzling it out: The current state of scientific knowledge on pre-kindergarten effects. A consensus statement*. Washington, DC: The Brookings Institution.

Rich, M. (2014, July 20). Obama to report widening of initiative for Black and Latino boys. *The New York Times*. Available from http://www.nytimes.com

Robinson, J. P., & Lubienski, S. T. (2011). The development of gender achievement gaps in mathematics and reading during elementary and middle school: Examining direct cognitive assessments and teacher ratings. *American Educational Research Journal*, *48*, 268-302.

Rutter, M., Caspi, A., Fergusson, D., Horwood, L. J., Goodman, R., Maughan, B., . . . Carroll, J. (2004). Sex differences in developmental reading disability: New findings from 4 epidemiological studies. *Journal of the American Medical Association*, *291*, 2007-2012.

Spelke, E. S. (2005). Sex differences in intrinsic aptitude for mathematics and science? A critical review. *American Psychologist*, *60*, 950-958. doi: 10.1037/0003-066X.60.9.950

Thomas, D. E., & Stevenson, H. (2009). Gender risks and education: The particular classroom challenges for urban low-income African American Boys. *Review of Research in Education*, *33*, 160-180. doi:10.3102/0091732X08327164

Sommers, C. H. (2013). *The war against boys: How misguided feminism is harming our young men*. New York, NY: Simon & Schuster.

Susman, E. J. (2006). Psychobiology of persistent antisocial behavior: Stress, early vulnerabilities and the attenuation hypothesis. *Neuroscience & Biobehavioral Reviews*, *30*, 376-389. doi:10.1016/j.neubiorev.2005.08.002

Votruba-Drzal, E., Coley, R. L., Maldonado-Carren, C., Li-Grining, C. P., & Chase-Lansdale, P. L. (2010). Child care and the development of behavior problems among economically disadvantaged children in middle childhood. *Child Development*, *81*, 1460-1474. doi:10.1111/j.1467-8624.2010.01485

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

Wood, D., Kaplan, R., & McLoyd, V. C. (2007). Gender differences in the educational expectations of urban, low-income African American youth: The role of parents and the school. *Journal of Youth and Adolescence*, *36*, 417-442.

Author Biographies

Clara G. Muschkin is an Associate Research Professor of public policy at the Duke University Sanford School of Public Policy, and Executive Director of the North Carolina Education Research Data Center. She is a Sociologist and Demographer with an interdisciplinary research focus. In her research, she asks how education policies that influence the composition and organization of educational institutions can influence student behavior and academic performance.

Helen F. Ladd is the Susan B. King Professor Emerita of Public Policy and Economics at Duke University's Sanford School of Public Policy. Her education research focuses on school finance and accountability, teacher labor markets, school choice, and early childhood programs. She is a past President of the Association for Public Policy Analysis and Management and a member of the National Academy of Education.

Kenneth A. Dodge is the Pritzker professor of Early Learning Policy Studies and professor of Psychology and Neuroscience at Duke University, and the founding and past director of the Center for Child and Family Policy. His research focuses on the development and prevention of aggressive and violent behaviors in children. He is a past president of the Society for Research in Child Development and a member of the National Academy of Medicine.

Yu Bai has a PhD in health policy and administration. Through his work as a Statistician at the Center for Child and Family Policy, he has contributed to numerous research projects that evaluate the effectiveness of education and health-related interventions for improving child outcomes.