ELIZABETH U. CASCIO
Dartmouth College

DIANE WHITMORE SCHANZENBACH
Northwestern University

The Impacts of Expanding Access to High-Quality Preschool Education

ABSTRACT President Obama's "Preschool for All" initiative calls for dramatic increases in the number of 4-year-olds enrolled in public preschool programs and in the quality of these programs nationwide. The preschools proposed by the initiative share many characteristics with the universal preschools that have been offered in Georgia and Oklahoma since the 1990s. This study draws together data from multiple sources to estimate the impacts of these "model" state programs on preschool enrollment and a broad set of family and child outcomes. We find that the state programs have increased the preschool enrollment rates of children from lower- and higher-income families alike. Among lower-income families, our findings also suggest that the programs have increased the amount of time mothers and children spend together on activities such as reading, the likelihood that mothers work, and children's test performance as late as eighth grade. Among higher-income families, however, we find that the programs have shifted children from private to public preschools, resulting in less of an impact on overall enrollment but a reduction in childcare expenses, and that they have had no positive effect on children's later test scores.

In his 2013 State of the Union address, President Obama proposed a sweeping reform of preschool education in the United States. His "Preschool for All" initiative calls for dramatic increases in the number of 4-year-olds in public preschool programs and improvements in the quality of these programs nationwide. The initiative would be funded by a $75 billion federal investment over 10 years, to be roughly matched by states, with federal dollars allocated to states based on the share of their 4-year-olds

who are from low- and moderate-income families.[1] Local school districts and other providers would be responsible for implementation, but in order to receive federal funding a state would have to adopt certain quality benchmarks related to early learning standards, teacher qualifications, and staffing ratios, as well as a plan for assessment.[2] The new preschools would be free for children from low- and moderate-income families but also accessible to children from higher-income families at a cost to be determined by individual states.

A large body of prior research suggests that early childhood education yields a high rate of return for children from low-income families. Prior studies of the long-run impacts of targeted preschool programs—for example, the Perry Preschool Project and the federally funded Head Start program—find that preschool produces not only private returns in the form of increased lifetime earnings but also public returns in the form of reductions in crime and in use of public assistance. Because of these externalities to preschool enrollment, economists would tend to agree that there is a role for public policy to expand preschool access.

The Obama plan builds on existing public preschool programs operated by state governments, which have greatly expanded over the past 30 years. Today, several states, including Tennessee, North Carolina, New Jersey, Washington, and Kentucky, have public programs that meet many of the quality benchmarks in the Obama plan, but they serve a small share of preschool-age students. Other states, including Wisconsin, Texas, and Florida, have programs that score high on access but low on quality. Very few states—most famously, Georgia and Oklahoma—have state-funded preschool programs that score high on both counts.

The Preschool for All proposal can be thought of as pushing all states toward the Georgia and Oklahoma model: an accessible preschool program with high standards. Existing evidence on the impacts of the Georgia and Oklahoma programs is focused on their short- to medium-term outcomes, and the findings are mixed. William Gormley and Ted Gayer (2005) find evidence that disadvantaged preschool attendees in Tulsa, Oklahoma, score

1. "Low- to moderate-income" is defined as income at or below 200 percent of the federal poverty line.

2. The Obama plan also includes expansions of Head Start for low-income 3-year-olds and Early Head Start for younger children, incentives for states to adopt full-day kindergarten classes, and extension of voluntary home visiting programs by nurses and social workers.

higher than their peers on tests at the end of the preschool year.[3] Maria Fitzpatrick (2008) finds that the Georgia program increases test scores of disadvantaged children as late as the fourth grade, but the effects appear smaller and less widespread than those from the Tulsa study. Fitzpatrick (2010) also finds no evidence that either of these state programs increases the labor supply of mothers of 4-year-olds, despite providing a 100 percent price subsidy for childcare on the extensive margin of employment.

To our knowledge, nothing is known to date about the impacts of these programs on child outcomes measured later than fourth grade or on measures of family well-being or behavior beyond maternal employment. This study addresses this gap in the literature, bringing together data from multiple sources to estimate the impacts of the Georgia and Oklahoma programs on a broad range of family and child outcomes. Our baseline empirical approach compares outcomes in these two "model" states to outcomes in the rest of the United States, before and after the introduction of their universal preschool programs, in a difference-in-differences (DD) framework. For outcomes where only more recent data are available, we take an alternative DD approach, comparing families with 5-year-olds and families with 4-year-olds, in Georgia and Oklahoma versus the rest of the United States. Where possible, we also combine the two strategies in a triple-difference framework, exploiting both the timing and age-targeting of the programs. Throughout, we present separate analyses by children's family background, since the alternatives to state programs might be weaker and the marginal returns to human capital investment higher for children from lower-income families. Instead of using family income directly, we present separate analyses by maternal education, which is a strong correlate of family income that is not directly affected by the implementation of these universal preschool programs and is observed for most children in our data sets.

We first show the stark difference in the effects of these programs on preschool enrollment patterns by maternal education. Among children whose mothers have no more than a high school degree, who are much less likely to be enrolled in preschool in the absence of a public option, we find that the Georgia and Oklahoma programs have sharply increased the likelihood of preschool enrollment at age 4. Our DD estimates imply that their preschool enrollment rates are now around 19 or 20 percentage points

3. Using a similar research design as Gormley and Gayer (2005), described in detail later in this paper, Wong and others (2008) also find evidence that the Oklahoma preschool program has improved children's end-of-preschool test performance.

higher than they otherwise would have been. On the other hand, the same empirical approach suggests that 4 or 5 out of every 10 public preschool enrollees whose mothers have at least some college education would otherwise have been enrolled in private preschools. As a result, the impact of these state programs on the overall preschool attendance of this latter group of children has been more muted, at an 11-to-14-percentage-point gain.

We then turn to examine the impact of the Georgia and Oklahoma programs on the behavior and well-being of families with 4-year-old children, continuing to split the data by maternal education. We hypothesize that the programs could affect how much time or money parents invest in their children, which could contribute to their academic achievement independently of preschool attendance itself. Our findings here are less precise, but align with those described above. The more highly educated families, for which private preschool enrollment falls significantly, see a significant reduction in their childcare spending. A back-of-the-envelope calculation based on our estimates suggests that this spending reduction amounts to an income transfer of between $3,000 and $5,600 for those families that switch out of private programs. By contrast, the less educated families, which gain more from the program in overall preschool participation, show larger declines in overall maternal time spent with children. However, this finding is counterbalanced by a positive impact the program has on the "quality" of time spent between mothers and children in these less educated families, such as time spent reading, playing, doing art projects, and talking.

In addition to its contribution to children's human capital formation, preschool serves as childcare. Indeed, another rationale for preschool expansions is the role that they would play as a childcare subsidy. An economic labor supply model frames a mother's decision to work as a function of her net wage, that is, the difference between her wage and the hourly price of childcare. Reducing the cost of preschool effectively increases a mother's net wage, making it more likely that she will participate in the labor force and, in turn, contribute to family income.[4] We find some evidence of an increase in the probability that less educated mothers are working when their children are 4 years old. However, the effect appears to be confined to the first few years after the program is in place, and estimates are relatively sensitive to changes in the specification.

4. Among mothers already participating in the labor force, the childcare subsidy will change the relative price of leisure, so the net impact on hours worked will depend on whether the income or substitution effect dominates.

Finally, we turn to these programs' reduced-form effects on children. We present separate analyses by a student's eligibility for free or reduced-price school lunch since maternal education is not reported in the underlying data. While the programs are still too young for us to estimate their truly long-term impacts, we are able to explore child outcomes as late as eighth grade with the benefit of more recent data than previously available.[5] For lower-income children, the evidence points to an impact of the programs on both reading and math scores in fourth grade, and while this effect diminishes substantially by eighth grade, it remains sizable in the case of math. However, for children from higher-income families, academic achievement does not appear to have improved on average in either grade, despite the fact that some of them were more likely to have attended preschool. While we cannot be completely certain of why this is, one possibility is that the programs induced a movement away from private preschools that were relatively high quality, leading to negative effects on the human capital of "switchers."

On the one hand, these findings would appear to suggest that an untargeted national preschool program would result in substantial substitution from private to public preschools, driving up costs and limiting program efficacy. On the other hand, the presence of higher-income children in the universal preschool classrooms in Georgia and Oklahoma—which may help to attract better teachers or have positive spillovers for lower-income children—may be what truly makes these programs "high quality." Regardless, our cost-benefit analysis in section V suggests that the benefits could outweigh the costs even with the substantial shifts from private to public preschool programs that we estimate.

I. Background

The Preschool for All initiative proposes a bold change in the role of the federal government in early childhood education. However, it builds on existing state efforts in preschool education and attempts to garner support from the large and well-identified literature on the long-term impacts of targeted preschool programs. In this section, we describe these state programs, elaborate on the small but growing literature on their impacts, and discuss key findings from the broader literature on preschool education.

5. The Georgia and Oklahoma preschool programs became universal in fall 1995 and fall 1998, respectively.

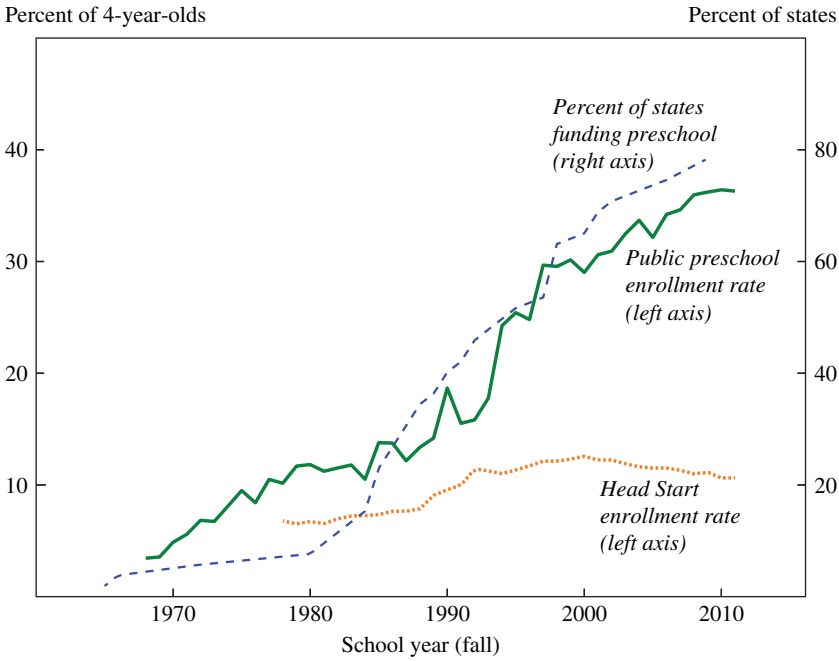
I.A. State-Funded Preschool Programs

Policy efforts at the state level have notably increased public preschool enrollment over the last 30 years. In 1980, only four states had subsidized the provision of preschool programs, and these programs were quite small.[6] But in the 5-year period from 1983 through 1987 alone, 11 states started their first preschool programs. Another eight states started programs over the next 5-year period, and by 2011 public preschool programs existed in 40 states and the District of Columbia. Figure 1 (right axis) plots the cumulative fraction of states that had funded preschool programs by year. Increases in the public preschool participation rates of 4-year-olds as measured in the October Current Population Survey (CPS) School Enrollment Supplements (left axis) track this state subsidization activity quite well, increasing by almost 25 percentage points between 1980 and 2011. Much of the public preschool enrollment of 4-year-olds prior to 1980 in the October CPS may be accounted for by Head Start, the already existing federal preschool program that targets children from low-income families. The Head Start enrollment rate of 4-year-olds (left axis) has risen little since the early 1990s.

There is significant variation across state preschool programs in who is eligible to attend, both in terms of age and in terms of targeting. Most programs admit only 4-year-olds, although approximately 13 percent of enrollment consists of 3-year-olds, according to surveys conducted by the National Institute for Early Education Research (NIEER) over the past decade. Most programs also target children from low-income families, although the income threshold for eligibility varies across states, and some state programs target children who have developmental delays or other risk factors regardless of income. Access is universal only in a handful of states, the longest-standing and most studied of which have been Georgia (universal access began in 1995) and Oklahoma (universal access began in 1998)—the two states that are also the focus of this study.

6. According to data compiled by the National Institute for Early Education Research (NIEER), these states were California (1965), New York (1966), Maryland (1980), and Oklahoma (1980). NIEER uses several criteria to identify state preschool programs. For example, the program has to be "funded, controlled, and directed by the state," serve preschool-age children, focus on early childhood education in a "group learning" environment, and "be distinct from the state's system for subsidized child care" (Barnett and others 2012; p. 21). As a result, the dates cited here may disagree with those reported elsewhere, such as by the Education Commission of the States.

Figure 1. Percent of 4-Year-Olds Enrolled in Public Preschool Programs and Percent of States Funding Preschool Programs, 1965–2011

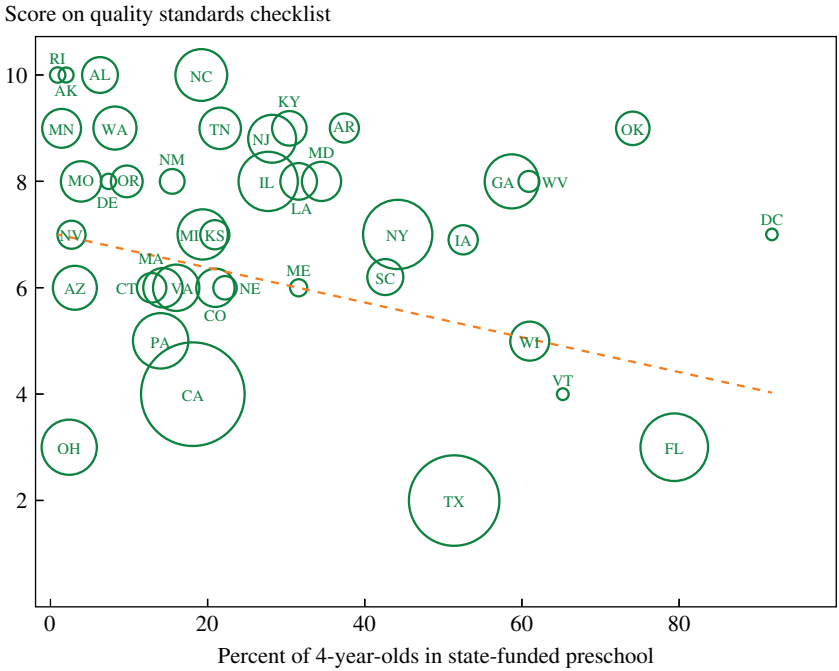


Sources: October CPS (public preschool enrollment rate), Head Start Bureau (numerator of Head Start enrollment rate), Vital Statistics (denominator of Head Start enrollment rate), and National Institute for Early Education Research (state funding dates).

The state programs differ not only in access but also in commonly used metrics of quality. For the past decade, NIEER has compiled state standards for preschool programs—related to curriculum, teacher education, class size, and support services—into an index with a maximum value of 10, giving each of 10 quality metrics equal weight; many of these metrics are incentivized by the Preschool for All initiative.[7] Figure 2 presents a scatterplot of this index against NIEER's internal estimate of the share

7. The 10 quality metrics are as follows: program has comprehensive early learning standards; teachers are required to have a BA; teachers are required to have specialized training in preschool education; assistant teachers are required to have a Child Development Associates (CDA) degree (or equivalent); at least 15 hours per year of in-service are required; the maximum class size is 20 students; staff-to-child ratios are 1-to-10 or better; program offers vision, hearing, health, and one support service; program offers at least one meal; program offers site visits.

Figure 2. Relationship between Quality and Access in State-Funded Preschool Programs, 2011–12 School Year[a]



Source: National Institute for Early Education Research (2012).

a. Dot sizes represent Census estimates of the state's 4-year-old population. The dashed line represents the regression fit, weighting by this figure; the unweighted fit is substantively similar. The quality standards checklist gives equal weight to each of 10 factors: program has comprehensive early learning standards; teachers are required to have a BA; teachers are required to have specialized training in preschool; assistant teachers required to have a Child Development Associates (CDA) degree (or equivalent); at least 15 hours per year of in-service required; the maximum class size is 20 students; staff to child ratios are 1-to-10 or better; program offers vision, hearing, health, and one support service; program offers at least one meal; program offers site visits.

of 4-year-olds in state-sponsored preschool programs as of the 2011–12 school year, the most recent with data available; dot sizes represent Census estimates of the state 4-year-old population, as reported by the NIEER. There appears to be a slight trade-off between access and quality according to this index, although the slope of the fitted line is not statistically significantly different from zero at conventional levels.[8] There is substantial

8. The slope of the fitted line (heteroskedasticity-robust standard error), weighting by the 4-year-old population, is –0.033 (0.025). Unweighted, the slope estimate is –0.023 (0.015).

variation in the index at all levels of access. For example, the programs in Georgia and Oklahoma, as well as those in West Virginia and the District of Columbia, are not only high-access but also high-quality, according to the index. However, other states with significant access, such as Florida, Texas, Vermont, and Wisconsin, score low on the index, and there is significant variation in this measure of quality across states where access is quite restricted.

Because our empirical analysis focuses on estimating the impacts of the Georgia and Oklahoma programs, it is useful to describe these programs in more detail.[9] In 2011–12, NIEER estimates that enrollment rates of 4-year-olds in state-funded preschool in Georgia and Oklahoma were 59 percent and 74 percent, respectively, with both programs serving that age group exclusively.[10] Both states also score high on the NIEER index, most recently receiving scores of 8 and 9, respectively (Barnett and others 2012). For example, both states require comprehensive learning standards, and both require that a classroom's lead teacher hold a bachelor's degree, be certified in early childhood education, and participate in annual in-service training.[11] Until recently, both states also mandated that the student-to-teacher ratio in preschool classrooms be no higher than 10-to-1 and that class sizes be no higher than 20.[12]

While these programs rank among the highest in the nation in access and quality according to NIEER, they differ from one another in several respects. Oklahoma's program is based in local school systems, and its funding runs from state governments to local school districts, which can choose to run half-day programs, full-day programs, or both. Spending per child is almost $7,427, with $3,652 coming from state sources and the remainder from a combination of local and federal contributions. By

9. We choose not to include West Virginia, since its program has only recently become high-quality on the NIEER scale. For example, while early learning standards were adopted in West Virginia in 2004–05, they have been in place in Georgia and Oklahoma since 1996–97 (Barnett and others 2012). Likewise, Washington, D.C.'s program only began in 2008–09. See online appendix A for discussion of how our findings change when Washington, D.C., and West Virginia are recoded as treatment states. (Online appendixes for papers in this volume may be found at the Brookings Papers website, www.brookings.edu/bpea, under "Past Editions.")

10. We do not use the NIEER estimates of enrollment rates in the analysis to follow, because the earliest available data are from the 2001–02 school year.

11. The Oklahoma program loses one point on the index because it does not require assistant teachers to have a CDA (or equivalent).

12. Georgia has recently lost two points on the index on this account, since it currently mandates a maximum class size of 22 and a maximum teacher-student ratio of 1-to-11. A recent state-commissioned evaluation of the Georgia program characterizes its classroom practices as of medium quality (Peisner-Feinberg, Schaaf, and LaForett 2013).

contrast, Georgia's program often runs through private childcare centers, and its total spending per child is only $3,490 and is drawn entirely from state sources. Since both models are possible under President Obama's proposed plan, we focus on estimating the average effect across the two programs throughout, giving each program equal weight.[13]

I.B. Previous Research on the Georgia and Oklahoma Programs

Given the policy relevance of the universal preschool programs in Georgia and Oklahoma, it is not surprising that they have already been topics of study. The primary area of interest has been children's academic achievement. In an evaluation of the Oklahoma program using data from the city of Tulsa, Gormley and Gayer (2005) leverage the sharp cutoff in eligibility to participate in the program based on a child's date of birth and find positive impacts of participation on cognitive measures at the end of the preschool year (see also Gormley, Phillips, and Gayer 2008). Impacts are largest for blacks and Hispanics and for low-income children, with little improvement for whites. Vivian Wong and others (2008) use the same research design but different data to estimate the effects of the Oklahoma program, also finding improvements in Peabody Picture Vocabulary Test scores at the end of the preschool year.

Less is known about outcomes in later grades. Fitzpatrick (2008) evaluates the medium-term impacts of Georgia's universal preschool program using fourth-grade outcomes measured in the National Assessment of Educational Progress (NAEP). Using a cross-state DD approach, she finds positive effects on fourth-grade NAEP test scores and the probability of being on-grade. The impacts are most consistently positive among disadvantaged non-urban students and more mixed among other demographic groups. They also appear smaller than the effects found in the Tulsa study, but this is not surprising given the pattern of "fadeout" in test score effects that pervades the early education evaluation literature. As described below, fadeout does not preclude positive impacts on a child's longer-term economic and social well-being.

These impacts on children's human capital are reduced-form, reflecting not only program participation itself but also potential changes in parental investments in children in response to the program. The literature on this

13. We discuss differences in some program impacts by state below.

mechanism is quite thin by comparison. Using a regression-discontinuity (RD) design similar to that used in the Tulsa studies, Fitzpatrick (2010) finds no statistically significant positive impacts of a child's eligibility for state-funded preschool on his or her mother's chances of working. Thus, most women appear to be infra-marginal with respect to the price subsidy for childcare implicit in these programs, that is, many women may switch from private preschools and other childcare arrangements to the state program without changing their employment status. As a result, the programs also do not appear to change family income, either by reducing public assistance receipt or increasing maternal earnings, when a child is 4 years old.[14]

We believe we are the first to leverage both the Georgia and Oklahoma expansions in the same DD framework to estimate their impacts on child and family outcomes.[15] There are some advantages to a DD approach over an RD one for the question at hand. One advantage is that the counterfactual to the universal program is captured by the experiences of 4-year-olds in other states after the initiatives are passed. In an RD approach, by contrast, the counterfactual is approximated by the experiences of children in the same state at a given point in time who will enter preschool the following year. To the extent that the alternatives would be changing over time—e.g., enrollment in private preschools might be increasing in control states—a DD approach using other states as controls might better represent what

14. A parallel literature exists on the impacts of universal kindergarten programs for 5-year-olds. Exploiting variation across states in the timing of first state funding for kindergarten in the 1960s and 1970s, Cascio (2009a) finds little evidence to suggest that kindergarten improved a child's long-term social and economic well-being, while Cascio (2009b) finds an impact on maternal labor supply, but only for single mothers with no children under the age of 5. On the other hand, attempting to leverage age eligibility by using quarter of birth as an instrument for kindergarten attendance in 1980, Gelbach (2002) finds more widespread increases in maternal employment in response to kindergarten, although Fitzpatrick (2012) provides evidence that estimates of maternal employment impacts using the quarter-of-birth instrument are quite different from those arrived at using an RD design based on exact date of birth and school entry cutoff birthdates. While interesting, these studies are arguably less relevant to the question at hand, given that the kindergarten programs under study were probably lower-quality than the current Georgia and Oklahoma preschool programs, that the counterfactual to kindergarten enrollment in the 1960s and 1970s and even 1980 was different than the counterfactual to preschool enrollment today, and that 5-year-olds are at a different point in development than 4-year-olds.

15. Thus, we focus on the demand side. In a recent study, Bassok, Fitzpatrick, and Loeb (2012) use a DD approach to estimate the effects of these particular programs on childcare providers.

would have occurred if the program had not been introduced.[16] In addition, using an RD approach one can only evaluate short-term child outcomes, like end-of-preschool test scores, or contemporaneous outcomes, like maternal employment when a child is 4 years old.[17]

A key contribution of the present study is thus to estimate these programs' impacts on child outcomes beyond the fourth grade test scores considered by Fitzpatrick (2008). Another contribution is to estimate the effects on contemporaneous, family-level outcomes beyond maternal labor supply. Much of what is known about the impacts of preschool programs on outcomes such as these is based on evidence from small-scale experiments in preschool education and the larger-scale federally funded Head Start program.

I.C. Previous Research on the Impacts of Targeted Preschool Programs

Arguably the most famous experimental preschool program was the Perry Preschool program, a 2-year intervention in the early 1960s involving half-day school attendance and weekly home visits for extremely disadvantaged 3- and 4-year-old African American children living in Ypsilanti, Michigan.[18] Students were randomly assigned to receive treatment or not, and follow-up data on the participants have been collected through age 40. Initial findings from Perry showed increases in IQ scores for the treated group; however, these initial gains did not persist, and by age 10 there was no measured difference in IQ scores between the treatment and comparison group (Gramlich 1986; Schweinhart and others 2005). Nonetheless, the Perry treatment students performed statistically significantly better in

16. Even state programs ranking lower on the ladders of access or quality have been found to have positive effects on children's outcomes, such as the low-quality, high-access program in Florida (Figlio and Roth 2009) and the high-quality, low-access programs in North Carolina (Ladd, Muschkin, and Dodge 2012) and Tennessee (Lipsey and others 2013a). The RD study by Wong and others (2008) also uncovers impacts of the preschool programs in Michigan, New Jersey, South Carolina, and West Virginia on print awareness at the end of the preschool year. In Tennessee, where the evaluation relied on a randomized control design, there were no longer measurable positive impacts on cognitive skills by the end of first grade, although there was some evidence of lasting gains on non-cognitive measures (Lipsey and others 2013b).

17. Another drawback of the RD approach is that children on either side of the age cutoffs have different ages relative to their classmates. To the extent that these peer effects influence outcomes, the estimated policy impact may be biased.

18. More specifically, the Perry program involved school attendance for 3 hours per day, 5 days a week from October through May, and weekly home visits for 90 minutes by a teacher to discuss a child's progress and instruct parents on how to provide an academically enriching environment at home (reading to children, counting with them, and so on).

school than the control group: they were absent fewer days and less likely to have been assigned to special education, had fewer failing grades and higher high school grade point averages, were more likely to graduate from high school, and generally reported more positive attitudes toward schooling. These improvements persisted into adulthood, by which time the treatment group was statistically significantly more likely to be employed and less likely either to have been arrested or to have received transfer payments such as cash welfare or food stamps.[19] When the improvements in long-term outcomes are monetized and discounted back to the start of the program, the benefits outweigh the costs by an estimated eight to one ratio (Heckman 2006).[20] Recent work by James Heckman, Rodrigo Pinto, and Peter Savelyev (2013) finds that the program induced changes in personality skills, which in turn can explain a large portion of the improvement in adult outcomes.

The high rate of return in the Perry program probably represents an upper bound on the return to any universal program for disadvantaged children today. While the Perry treatment was of high quality in many of the same respects as the Georgia and Oklahoma programs, the alternatives to participating in any program for lower-income children have expanded dramatically over time. Today, not only can lower-income children attend Head Start; they have access to some targeted state funded programs, as described above, and some find themselves in center-based or informal childcare arrangements while their mothers work. The marginal benefit of attending a high-quality program for low-income children today would therefore likely be smaller than it was in the 1960s. The participants in Perry were also extraordinarily disadvantaged, and marginal investments in their human capital might have had high returns.

Based in part on the early successes of Perry, Head Start was started in 1965 as part of the "War on Poverty." Head Start is a large public preschool program for low-income children that is generally thought to be lower in quality than the Perry program yet higher in quality than the childcare that is typically available to low-income parents (Currie 2001). To put relative quality into context, Head Start has been estimated to cost more than many

19. A similar pattern of findings has emerged for other preschool experiments, for example North Carolina's Abecedarian Project (Masse and Barnett 2002; Anderson 2008).

20. Recent reanalysis using different methods to account for multiple hypothesis testing by Anderson (2008) generally confirms these findings, but finds that the long-term benefits in Perry are concentrated among female participants. Heckman and others (2010) calculate rates of return separately by gender, and find that returns are significantly different from zero for both males and females.

state-funded preschool programs but less than Perry and other high-quality preschool programs (Currie 2001). However, Head Start ranks lower on the NIEER scale than many state-funded preschool programs, averaging a score just below 5 (Espinosa 2002).[21] The program also reaches further up the income distribution than Perry did, and estimates of its longer-term impacts have been drawn not only from the earliest cohorts to attend, for whom the alternatives were relatively limited (Garces, Thomas, and Currie 2002), but also from among more recent cohorts with a wider range of alternatives (Deming 2009).

The first randomized evaluation of Head Start, the Head Start Impact Study, was conducted in 2002. Head Start participants in that study saw faster improvements in language and literacy skills than their non–Head Start counterparts. However, these improvements dissipated quickly, and there were no measurable differences in the test scores of Head Start and control children as of third grade (Puma and others 2012). Unfortunately, the Head Start Impact Study was undertaken too recently to know whether children randomized into Head Start will have better non-test outcomes over the longer term.

The quasi-experimental literature to date on the longer-term impacts of Head Start is promising, however. The predominant research design in this literature has been to compare outcomes across siblings who were exposed to different preschool environments (such as Head Start vs. another preschool or no preschool). Much of the within-family variation appears to be idiosyncratic and may be driven by the availability of slots at local programs, which are often oversubscribed.[22] Findings from sibling comparisons suggest that Head Start participation, relative to attending either no preschool or a non–Head Start preschool, has a substantial positive effect on vocabulary test scores during elementary school and makes a child less likely to repeat a grade (Currie and Thomas 1995; Deming 2009). While test score gains fade to a fraction of their initial levels by ages 11 to 14 (and are no longer statistically different from zero), during their teen

21. For example, Head Start teachers tend to have relatively weak academic credentials. However, the quality of Head Start may have improved recently in response to recent policy initiatives. Currently, 56 percent of Head Start teachers hold a bachelor's degree, and another 30 percent hold an associate's degree, up from 30 and 16 percent, respectively, in 1999. Over the same time period, the average child-to-teacher ratio dropped from 19.8 to 17.6, and the average staff turnover rate dropped from 16.8 to 13.2 percent (Bassok 2012).

22. On the other hand, to the extent that a change in a family's economic circumstances renders one sibling eligible and another ineligible, this approach could understate the impact of the program.

years Head Start participants are nevertheless less likely to have ever been charged with a crime or to have become a parent and are more likely to complete high school and attend college as young adults (Deming 2009; Garces, Thomas, and Currie 2002).[23]

The pattern of findings in both the Perry program and the quasi-experimental Head Start literature—strong contemporaneous impacts, followed by smaller medium-term impacts, culminating in strong adult impacts—has been documented repeatedly in evaluations of early childhood interventions.[24] For this reason, it would be ideal to measure impacts on longer-term child outcomes besides test scores. Unfortunately, the Georgia and Oklahoma programs were introduced too recently for us to estimate their impacts on adult outcomes. However, in our data we already see the beginning of the same pattern of program impacts on test scores that fade over time.

By comparison to the literature on child outcomes, the literature on the impacts of Perry and Head Start on families is quite thin. Using data and random variation in Head Start participation from the 2002 Head Start Impact Study, Alexander Gelber and Adam Isen (2013) find that Head Start induces parents to become more involved in their children's learning, as measured by increases in time spent reading to kids, in the number of visits to cultural events, and in time spent with nonresident fathers. These improvements in the home environment persist even after Head Start ends.

II. The Preschool Experience

For the Preschool for All initiative to have an impact on children's human capital, it must first affect their preschool enrollment. The first-order question for our analysis is therefore this: How have the state programs in Georgia and Oklahoma affected children's preschool experiences?

23. Ludwig and Miller (2007) also find evidence that Head Start participation increases educational attainment, exploiting the sharp difference in special grant-writing assistance afforded to counties with similar poverty rates at the program's inception.

24. James Heckman and his coauthors (such as Heckman 2006; Knudsen and others 2006) have put forth a helpful framework to explain this pattern, relying on a distinction between cognitive and non-cognitive skills. While cognitive skills encompass the ability to add, subtract, read, and so on, and are readily measured on traditional standardized tests taken by school-age children, so-called "non-cognitive" skills are less likely to be measured on achievement tests and include social, behavioral, and emotional skills, such as persistence, behavior, and the ability to cooperate. Both types of skills are important inputs to adult outcomes like employment, wage levels, marriage, and criminal activity. The hypothesized mechanism through which long-term gains are realized is through a permanent improvement in non-cognitive skills that persists after the temporary gains in cognitive skills have faded.

It stands to reason that these programs have unambiguously increased the chances that disadvantaged children attend preschool and the average quality of preschools attended, given the constraints faced by lower-income families. However, a substantial fraction of these programs' budgets could be subsidizing higher-income families, whose children may already have a relatively high probability of attending preschool and for whom program quality may already be quite high. For these children, there is not only less scope to increase preschool enrollment, but the private programs "crowded-out" may also be at least as high in quality as the state program, opening the possibility of negative impacts on human capital.[25]

Our first analytical challenge is therefore to estimate how these programs have affected preschool experiences and how this has varied by a child's family background. The best available data for this purpose are from the October CPS School Enrollment Supplements. The October supplements provide annual data back to 1968 (with complete information on state of residence dating to 1977) on the preschool enrollment of 4-year-olds, both overall and by school type, public or private.[26] We use the school type question to provide insight into substitution from private to public preschools, noting that the possible inclination of survey respondents to classify state-funded preschool as private enrollment in Georgia, where the program operates in part through private centers, should bias us against finding evidence in favor of it.

Importantly, because entire households are surveyed, the October supplements allow us to match children to other family members and so obtain information on a family's current socioeconomic status (SES). While the Preschool for All initiative conceives of SES as family income, family income itself would be an outcome of the program if it were to change maternal employment—a question that we address below. We therefore think a more desirable sample split is based on a strong predetermined *predictor* of family income. In this spirit, we split the sample into two

25. Of course, this is not to say that choice is welfare-reducing, since the reduction in human capital will be offset by the family's lower spending on childcare. We explore the impacts of these programs on consumption of childcare services in the next section.

26. Ideally, we would have detailed enough information to classify children into "school entry cohorts" based on their birthdays and the minimum age of kindergarten entry in their state. While information on school entry regulations is available, we only know a child's age as of October. Fortunately, the vast majority of children who are 4 years old in October would be eligible for preschool rather than kindergarten. In Georgia and Oklahoma, for instance, by September 1 children must be 4 years old to enter preschool and 5 years old to enter kindergarten.

maternal education categories: children of mothers with no more than a high school degree and children of mothers with at least some college education.[27] Results are substantively similar if we instead stratify by paternal education or whether a family's income-to-poverty ratio is above or below 185 percent, the cutoff for reduced-price lunch (see online appendix B).

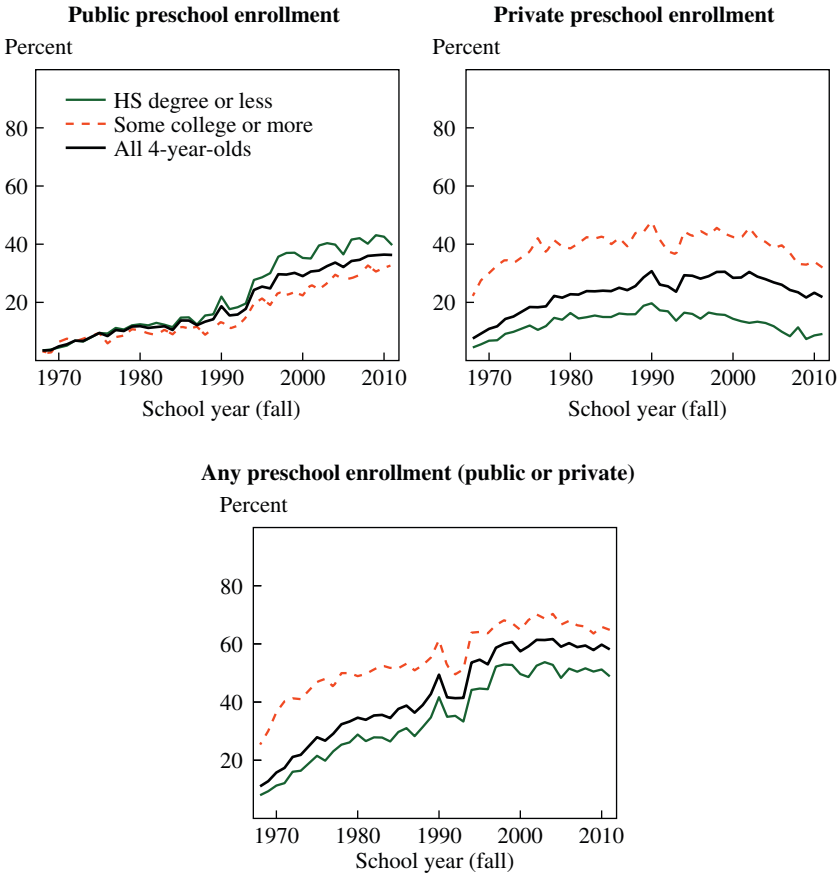
II.A. National Trends in Preschool Enrollment

To set the stage, figure 3 shows national trends in public preschool enrollment, private preschool enrollment, and enrollment in any preschool (the sum of public and private) from 1968 through 2011, both for the population of 4-year-olds overall and for each of the two maternal education categories. As noted in reference to figure 1, the overall national trend in public preschool enrollment (solid black line in upper-left panel) lines up well with the introduction of state funding for preschool. Given the targeted nature of most of these programs, it is not surprising that the increases in public preschool enrollment have been greater among the children of women with no more than a high school degree than among the children of women with some college or more. That said, by the end of the sample period the children of more educated women were still more likely to be attending any preschool, public or private (lower panel), continuing their long tradition of being more likely to attend a private program (upper-right panel).

But more striking is the fact that the private preschool enrollment rates of 4-year-olds declined markedly over roughly the same period that state-funded preschool programs expanded. Regardless of maternal education, private preschool enrollment rates in 2011 were not much higher than they had been in 1968, having peaked right before the recession in the early 1990s. Table 1 provides some suggestive evidence of the extent of substitution between public and private programs. Between 1985 and 2010, the private preschool enrollment rate of 4-year-olds whose mothers have at least some college education fell by 8.6 percentage points; among 4-year-olds whose mothers have no more than a high school degree, the decline in the private enrollment rate was almost as large, at 7 percentage points. The

27. We have also split the sample into four maternal education categories (high school dropout, high school degree only, some college only, and college degree or more) and have found that dividing the sample into two groups was a parsimonious way of capturing the patterns in the data. While the college-educated share of mothers has increased dramatically over the sample period (see, for example, Goldin, Katz, and Kuziemko 2006), maternal education did not change in response to the Georgia and Oklahoma programs (see online appendix table B-1).

Figure 3. Preschool Enrollment Rates of 4-Year-Olds, Overall and by Maternal Education, 1968–2011



Source: October CPS school enrollment supplements, 1968–2011.

declines in private enrollment among children in these groups represent significant shares—41 and 25 percent, respectively—of their increases in public preschool enrollment over the same period. As a result, the groups' gains in overall enrollment between 1985 and 2010 have been much more limited than the expansion of public enrollment alone would suggest: gains of 12.6 percentage points and 21.1 percentage points, respectively.

Although suggestive of crowd-out, the declines in private preschool enrollment over the past 25 years might have been caused by other factors, such as the stagnation in maternal labor force participation over the same

Table 1. Preschool Enrollment of 4-Year-Olds and Employment of Mothers of 4-Year-Olds, by Maternal Education: 1970, 1985, and 2010

	Year			*Change*	
	1970	*1985*	*2010*	*1970–85*	*1985–2010*
Mother has high school degree or less					
Enrolled in public preschool	4.4	13.7	41.7	9.2	28.1
Enrolled in private preschool	6.4	15.4	8.4	8.9	–7.0
Enrolled in any preschool	10.9	29.0	50.2	18.1	21.1
Mother employed last week	28.6	42.4	43.1	13.7	0.8
Mother has some college or more					
Enrolled in public preschool	5.5	10.6	31.8	5.1	21.2
Enrolled in private preschool	30.1	41.6	33.0	11.4	–8.6
Enrolled in any preschool	35.6	52.2	64.8	16.6	12.6
Mother employed last week	31.7	56.1	61.6	24.3	5.5

Source: Authors' calculations from the October CPS, 1969–71 (for 1970), 1984–86 (for 1985), and 2009–11 (for 2010).

period.[28] Indeed, the increases in private enrollment through 1985 occurred over a period of rising maternal employment rates (see table 1), suggesting a link between the two.[29] While this too is only a correlation in two national time series, it reminds us that private preschools also provide childcare, the demand for which is driven by factors beyond the availability of public alternatives.

We now turn to our formal estimates of the preschool enrollment responses to the Georgia and Oklahoma programs, using an empirical strategy that has the potential to account for those other factors.

II.B. Impacts of the Georgia and Oklahoma Programs on Preschool Enrollment

EMPIRICAL STRATEGY Our analysis separates out the trends in preschool enrollment in Georgia and Oklahoma from those for states elsewhere in the country, using the October CPS supplements from 1977 forward. With these data in hand, we can compare the changes in preschool enrollment in Georgia and Oklahoma after the introduction of their universal preschool

28. Increases since the mid-1980s in the state-mandated age at which children are eligible to enter kindergarten and in the practice of "red-shirting" (delaying a child's entrance into school) (Deming and Dynarski 2008) are other potential contributors to the downward trend in private preschool enrollment of 4-year-olds shown in figure 3.

29. A full-time series of maternal employment rates is provided in figure 5, top panel.

initiatives to changes in preschool enrollment in the rest of the country over the same period. Thus, instead of attempting to measure and control for all of the other factors affecting private preschool enrollment, we assume that the effects of these other factors are accurately embodied in what happened to preschool enrollment in other states.[30]

In its simplest incarnation, this quasi-experimental DD approach is captured in the following equation:

$$y_{st} = \theta\, post_{st} + \gamma_s + \delta_t + \varepsilon_{st}, \tag{1}$$

where y_{st} is a preschool enrollment rate of 4-year-olds in state s in October of year t (in percent terms); $post_{st}$ is an indicator variable set to one in Georgia from 1995 forward and in Oklahoma from 1998 forward, and set to zero otherwise; and γ_s and δ_t are vectors of state and year fixed effects, respectively. The state fixed effects account for fixed differences in preschool enrollment rates across states, while the year fixed effects account for common shocks to the preschool enrollment rate (that is, shocks to private preschool demand stemming from the business cycle). In some specifications, we also add a vector of state-by-year-varying controls, including child demographics and the state unemployment rate. ε_{st} is an error term, which represents unobserved determinants of enrollment rates.

Estimation of this simple model will identify the coefficient of interest, θ, only if none of the unobservables is correlated with the adoption of universal preschool. This assumption would be violated if, for example, the private preschool enrollment rates in Georgia and Oklahoma would have been on a steeper downward trajectory than elsewhere even without starting universal preschool programs. In this case, the estimates would imply more crowd-out as a result of universal preschool than has actually taken place. The model is also restrictive in the sense that it assumes that the effects of universal preschool on enrollment rates were both immediate and persistent.

We therefore begin by estimating a less restrictive event-study model, which allows us to test whether the model states were on different enrollment trajectories prior to introducing their programs, as well as whether the

30. Using other states in the southern census region to proxy for these other influences, our DD point estimates for enrollment and many of the other outcomes under study tend to be diminished in magnitude, but the general pattern of findings still holds (see online appendix C). The decline in the magnitude of effects is unsurprising, given that other southern states have made substantial investments in preschool, as suggested by figure 2.

impacts of these programs have remained constant as they have matured. In this model, we replace the $post_{st}$ indicator in equation 1 with a series of indicator variables for year relative to the year that universal preschool was introduced. Instead of creating an indicator for each individual year relative to the initiative, we create dummies for 3-year bins to reduce noise. So that the coefficients are identified, we omit the dummy for the three years immediately prior to the initiative (representing 1992–94 in Georgia and 1995–97 in Oklahoma).[31] This model provides us with transparent graphical evidence on the credibility of equation 1's identifying assumptions.

FINDINGS Figure 4 plots the event-study estimates for each of the three preschool enrollment rates. To facilitate the comparison across family background, we present the coefficient estimates for a given variable for children from both maternal education categories in the same graph, offsetting each series to the right and left of the midpoint of the relevant interval for ease of viewing. The capped lines around the coefficient estimates represent 90 percent confidence intervals.[32]

For children whose mothers have a high school degree or less (solid circles), the introduction of the Georgia and Oklahoma programs appears to have increased enrollment in public preschools by 15 to 20 percentage points relative to expectations based on public enrollment trends elsewhere in the country (upper-left panel). The point estimates are somewhat higher among children whose mothers have more education (open circles), centered on a 25- to 30-percentage-point impact. For both groups of children, each of the post-program coefficient estimates is statistically significant, and there is little evidence to suggest that the effects of the program on public enrollment have systematically grown or shrunk over time.[33]

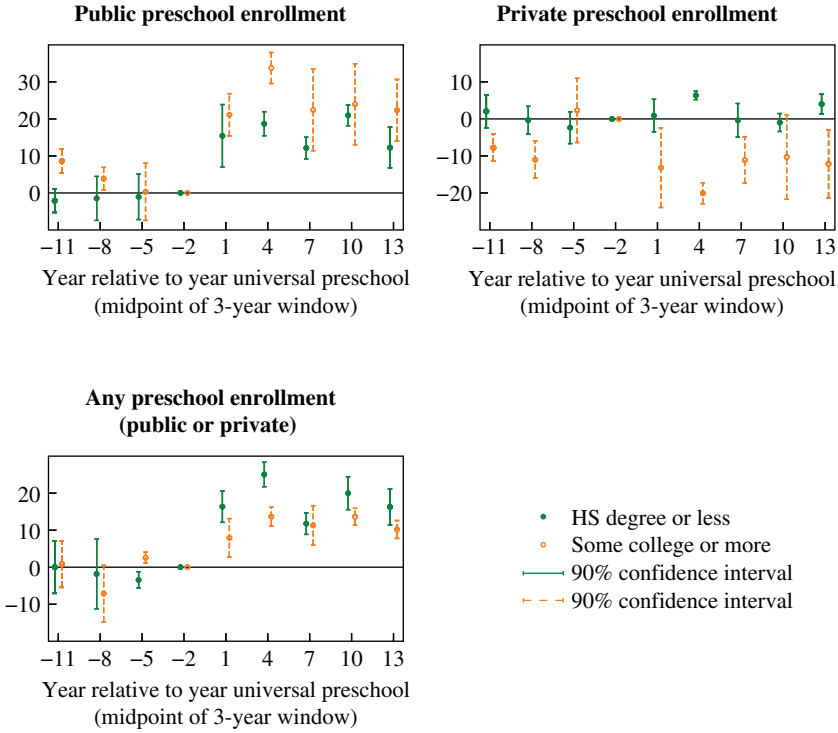
The remainder of the figure then provides evidence of differences by family background in the extent of substitution between public and private preschool in response to the initiatives. Enrollment in private preschool of children with less-educated mothers did not decline faster than elsewhere

31. The first and last indicators represent all prior and subsequent years, respectively.

32. Regressions are weighted by the number of observations in the October CPS micro data used to calculate the enrollment rate, and standard errors allow for heteroskedasticity and correlation in the error terms within states over time. We discuss weighting and calculation of standard errors in more detail below in reference to table 2.

33. This might seem somewhat surprising, given that administrative data suggest that universal preschool enrollment rates in Oklahoma increased somewhat gradually over time, almost doubling between the 1998–99 and 2005–06 school years (Bassok, Fitzpatrick, and Loeb 2012). The program was phased in much more rapidly in Georgia, however, and, as suggested by figure 3 (upper-left panel), the counterfactual is one of increasing public preschool enrollment.

Figure 4. Event-Study Estimates of the Effect of Introducing High-Quality Universal Preschool on Preschool Enrollment Rates of 4-Year-Olds, by Maternal Education[a]



Source: October CPS school enrollment supplements, 1977–2011.

a. States with high-quality preschool programs are Georgia (introduced in fall 1995) and Oklahoma (introduced in fall 1998). All regressions include state fixed effects and year fixed effects in addition to dummies for 3-year groupings of year relative to the year that high-quality preschool was introduced. The coefficients plotted at –11 represent 10 years or more prior to introduction, while the coefficients plotted at 13 represent 12 or more years after; the dummy at –2, representing the three years immediately prior to introduction, is omitted to identify the model. All regressions were weighted by the number of children used to calculate the enrollment rate, and standard errors were clustered on state.

in the country after universal preschool programs were introduced (upper-right panel). As a result, the increase in overall enrollment of these children tracks almost one-for-one with the increase in public enrollment (lower panel). On the other hand, for children with more highly educated mothers, private preschool enrollment declined between 10 and 20 percentage points more than expected based on national trends (upper-right panel), greatly diminishing the overall enrollment impact (lower panel) relative to program take-up (upper-left panel). On net, both groups of children were

Table 2. Difference-in-Differences Estimates of the Impact of High-Quality Universal Preschool on Preschool Enrollment Rates of 4-Year-Olds, by Maternal Education[a]

	Effect sizes		
Dependent variable (percent)	*Baseline DD 2-1*	*Add demographics 2-2*	*Add state linear trends 2-3*
Mother has high school degree or less (N = 1,785)			
Public preschool enrollment rate	17.49***†	16.99***†	17.13***†
	(1.326)	(1.342)	(1.136)
Private preschool enrollment rate	1.191	1.822	3.035***
	(2.349)	(1.974)	(0.851)
Effect of public on private	*0.068*	*0.107*	*0.177****
	(0.135)	*(0.116)*	*(0.0508)*
Overall preschool enrollment rate	18.68***	18.81***	20.17***†
	(2.487)	(2.434)	(1.435)
Mother has some college or more (N = 1,784)			
Public preschool enrollment rate	19.63***††	19.72***††	28.66***††
	(3.449)	(3.535)	(2.766)
Private preschool enrollment rate	–8.296††	–8.044††	–14.90***††
	(5.442)	(6.105)	(3.244)
Effect of public on private	*–0.423***	*–0.408**	*–0.520****
	(0.206)	*(0.239)*	*(0.0681)*
Overall preschool enrollment rate	11.34***	11.67***	13.77***
	(2.233)	(2.764)	(1.134)
Controls:			
State fixed effects	Y	Y	Y
Year fixed effects	Y	Y	Y
Nonwhite (percent), female (percent)	N	Y	Y
State unemployment rate	N	Y	Y
Linear trends for GA and OK	N	N	Y

Source: Authors' regressions based on state-by-year-by-age averages of microdata from the October CPS School Enrollment Supplements, 1977–2011.

a. Each non-italicized coefficient is a separate OLS estimate of θ from model (1). See the text for details. Each italicized coefficient is the two-stage least squares estimate of the effect of the public preschool enrollment rate on the private preschool enrollment rate, where the instrument for the public preschool enrollment rate is the *post* indicator. Regressions are weighted by the number of children used to calculate the enrollment rate. Standard errors clustered on state are in parentheses. Asterisks indicate statistical significance based on these conventional standard errors at the ***1 percent, **5 percent, or *10 percent level. Daggers indicate statistical significance based on Conley and Taber (2011) confidence intervals at the †† 5 percent or † 10 percent level (and apply to non-italicized DD estimates only).

more likely to be enrolled in preschool after the introduction of universal preschool, but the overall enrollment effects appear to be smaller for higher-SES children.

Table 2 presents the DD estimates, helping us to put more concrete numbers on these patterns. Because we are working with state-by-year level means of the October CPS micro data, we weight the regressions (as we

weighted the regression estimates underlying figure 4) by the number of children used to calculate the enrollment rates; unweighted estimates are similar (see online appendix table D-1). We also present standard errors that are clustered on state. Because these standard errors rely on large-sample approximations that do not apply in our application, we also calculate confidence intervals for the DD estimates that are robust to having only two treatment states, using the baseline method described in Conley and Taber (2011). We denote statistical significance based on this method with dagger superscripts in table 2 and present the confidence intervals themselves in online appendix table E-1.[34]

As anticipated from the graphs, estimates of the simple DD model in equation 1 (column 2-1) yield larger impacts of universal preschool on the public preschool enrollment of children of more-educated mothers—a 19.6-percentage-point increase—versus a 17.5-percentage-point increase for the children of less-educated mothers. Controlling for state unemployment rates and racial and gender composition does not change the estimates much (column 2-2), though the difference across family background is made more stark when we add separate linear time trends for Georgia and Oklahoma in an attempt to account for the relatively strong downward trend in public enrollment among the children of more-educated moms in these states prior to universal preschool availability (column 2-3).[35]

However, the implications regarding substitution between private and public preschool are the same regardless of specification. Among children whose mothers have at least some college education, the reductions in private enrollment following the introduction of public preschool—the effect of public on private enrollment, shown in italics in the table—imply that the likelihood of switching is 41 to 52 percent, depending on the

34. The remainder of online appendixes D and E show and discuss, respectively, unweighted estimation and Conley and Taber (2011) confidence intervals for later tables in the paper where these methods are applicable. These confidence intervals are probably conservative, leading us to reject the null hypothesis too infrequently (for example, using the 95 percent confidence interval, we would reject the null less than 5 percent of the time).

35. One concern with a specification such as this, explored most prominently by Wolfers (2006), is that when a policy's effects grow or shrink over time, state-specific trends will pick up these dynamics in addition to pre-existing trends, potentially generating substantial bias in DD coefficients. When we follow his suggestion of including the series of post-initiative indicators from our event-study model in lieu of $post_{st}$ in equation (1), our substantive conclusions are unchanged.

specification.[36] That is, for every 10 children with more-educated mothers who enroll in public preschool, 4 or 5 would otherwise have been enrolled in private school. This figure lines up very well with the back-of-the-envelope calculations based on the national trends in table 1. Unlike these simple calculations, however, we find no evidence of a shift from private to public preschool among children whose mothers have a high school degree or less. In fact, for this subpopulation, the evidence suggests that private preschool enrollment rates *rose* in the post-initiative period. This could reflect misreporting by households in Georgia, where some state-funded preschools are operated in private childcare centers. Indeed, when we allow for separate impacts by state, we detect a significant increase in the private preschool enrollment rates among the children of less educated mothers only in Georgia (see online appendix table F-1).[37]

The bottom line is that the high-quality universal preschool programs in Georgia and Oklahoma increased the likelihood of enrolling in preschool at age four by a greater amount among lower-SES children (increasing by 18.7 to 20.2 percentage points) than among higher-SES children (increasing by 11.3 to 13.8 points).

III. Family Well-Being and Behavior

Although a shift in enrollment from private to public preschool programs might diminish the potential impact of universal preschool programs on the human capital of higher-SES children, it could make some higher-SES families better off in the short term by reducing their childcare expenses. By increasing a mother's (or primary caregiver's) wage net of childcare costs for at least part of the work day, the childcare price subsidy implicit in universal preschool may also increase labor force participation—and thus a family's income and child investment possibilities—regardless of family background. Further, while these programs might reduce the amount

36. These estimates were obtained by regressing the private preschool enrollment rate on the public preschool enrollment rate, instrumenting for the latter with the *post* indicator, and weighting by the number of children used to construct the private enrollment rate. The two-stage least squares estimator is thus the Wald estimator—the ratio of the DD estimates for private and public enrollment rates.

37. We also find that Georgia's program led to significantly less reduction in private preschool enrollment among children with more-educated mothers than Oklahoma's program. On net, the impacts of universal preschool on the likelihood of being enrolled in any preschool, public or private, were greater in Georgia than in Oklahoma, albeit statistically significantly so only in specifications 2-1 and 2-2 in table 2 and using conventional methods of inference (see online appendix table F-1).

of time that many 4-year-olds will spend at home, they could also change the quality of the remaining time that parents spend with their children, depending on whether parents view these investments as substitutes for or complements to formal education.

Each of these potential responses to the program—the allocation of parental time, consumption patterns, and labor supply—might affect a child's human capital accumulation independently of preschool itself. In this sense, any impacts of universal preschool on children's test scores will be reduced-form, reflecting not only participation in the preschool program itself, but also changes in parental investments. In this section, we attempt to understand the magnitude of these changes before turning to our estimates for children's test scores.

III.A. Childcare Spending

The movement of children from private preschool programs into state-funded ones should put money in a family's pocketbook. But how much do families thereby save in childcare expenses? In other words, how large an income transfer does universal preschool provide? To address this question, we turn to the Consumer Expenditure Survey (CEX), which has asked comparable questions about expenditures on childcare and daycare services since 1994. Note that this data set aggregates all household childcare expenses, so it does not allow us to isolate spending on just the 4- or 5-year-old child. Limited pre-initiative data are available, and the number of households with 4-year-olds is much smaller than in the October CPS. As a result, our estimates using these data are less precise and rely on stronger identifying assumptions—and our conclusions are thus more tentative—than those for enrollment presented above.

In light of the data limitations, we take an alternative DD approach, examining the difference in childcare expenses between families with 4-year-olds and families with 5-year-olds and seeing how that difference, in Georgia and Oklahoma, compares with the same difference elsewhere in the country.[38] We restrict the sample to the period after the model preschool

38. An alternative comparison group would be families with 3-year-olds. We prefer families with 5-year-olds as a comparison group, because they are less likely than families with 3-year-olds to have been affected by state preschool funding initiatives (e.g., through their eligibility for preschool programs in the comparison states). Our estimates of the impact of universal preschool on childcare expenses are smaller in magnitude when we use families with 3-year-olds as a comparison group (see online appendix table G-1).

programs are in place, analyzing monthly spending from September 1998 through June 2011.[39] The equation of interest is:

$$(2) \qquad ccspend_{as} = \alpha + \theta age4_a treat_s + \delta age4_a + \gamma treat_s + \varepsilon_{as},$$

where $ccspend_{as}$ represents average monthly spending (in inflation-adjusted 2012 dollars) on childcare for families in state s with at least one child age a; $age4_a$ is an indicator variable set to one if that average corresponds to families with a 4-year-old; and $treat_s$ is an indicator variable set to one if the state in question is Georgia or Oklahoma.[40]

The coefficient of interest in equation 2 is on the interaction between these two indicators. It captures the extent to which universal preschool programs reduce the gap in childcare spending between families with 4-year-olds and families with 5-year-olds, all of whom have access to at least half-day kindergarten. This coefficient will capture the effect of universal preschool if there is no other reason to expect a smaller gap in the model states. For example, the childcare spending gap would be smaller if Georgia and Oklahoma were less likely to offer full-day kindergarten for 5-year-olds than other states. In fact, the opposite appears to be the case.[41]

The first column (3-1) of table 3 presents estimates of equation 2, again splitting the data into the two maternal education groups. Consider first the estimates for families with college-educated mothers (lower panel). The coefficient of interest is a statistically significant $66 reduction in monthly childcare expenses (in real 2012 dollars). This almost perfectly offsets the

39. We drop spending observations for July and August, when preschool is typically not in session. The CEX suppresses state of residence for confidentiality purposes in some cases. For example, starting in 2006, we can no longer identify Oklahoma in the data because it is suppressed. Because the data are not evenly dispersed over time for all states, we also collapse the monthly data to the state-by-year level, separately by age and mother's education group, and include year fixed effects in the models.

40. As in our estimates based on the October CPS, we weight by the number of observations used to calculate the state-by-age group-by-year averages and present standard errors clustered on states in parentheses. The unweighted estimates are actually much more stable across the alternative measures of childcare expenses (see online appendix table D-2). However, calculated using the Conley and Taber approach, 90 percent confidence intervals on all of the DD coefficients in table 3 include zero (see online appendix table E-2).

41. Using the October supplements from 1998 through 2011 collapsed to state-by-year level averages and weighting by cell size, we find that 52.9 percent of 5-year-olds in Georgia and Oklahoma whose mothers have at most a high school degree attend full-day kindergarten; elsewhere in the country, this figure is 46.9 percent. For the 5-year-old children of women with at least some college education, full-day kindergarten enrollment rates are 51.8 and 48.8 percent for the treatment and comparison states, respectively.

Table 3. Difference-in-Differences Estimates of the Impact of High-Quality Universal Preschool on Monthly Spending on Child Care, by Maternal Education[a]

Coefficient on:	*Baseline DD* 3-1	*Topcode spending*[b] 3-2	*Use median spending*[c] 3-3
	Mother has high school degree or less (N = 920)		
Age 4 × GA or OK	35.48***	34.94***	–1.723
	(9.428)	(9.415)	(4.145)
Age 4	7.747	8.414	0.0475
	(5.716)	(5.703)	(3.582)
GA or OK	–31.34***	–30.36***	–9.438
	(5.595)	(5.445)	(6.495)
	Mother has some college or more (N = 1,008)		
Age 4 × GA or OK	–65.86***	–49.69**	–56.17***
	(21.66)	(19.45)	(14.12)
Age 4	70.37***	63.29***	54.39***
	(14.25)	(12.40)	(11.57)
GA or OK	–32.30**	–31.47***	4.046
	(12.87)	(10.94)	(13.90)
Additional controls:			
Year fixed effects	Y	Y	Y

Source: Authors' regressions based on state-by-year-by presence of 4-year-old averages of microdata from the Consumer Expenditure Survey, September 1998–June 2011.

a. Each panel and column represents a separate estimate of model (2). Average monthly childcare spending is in real 2012 dollars. Sample is limited to families with a 4- and/or 5-year-old child. Regressions are weighted by the number of families used to calculate the dependent variable. Standard errors clustered on state are in parentheses. Asterisks indicate statistical significance at the ***1 percent or **5 percent level.

b. In column 3-2, we topcode spending at the 99th percentile before collapsing the data to state-by-year-by-age means.

c. In column 3-3, we collapse the data to cell medians instead of cell means.

additional average childcare spending for families with 4-year-olds versus 5-year-olds (the coefficient on age_a), which is estimated to be $70 per month. Because the sample sizes are small and we are concerned about outliers, in column 3-2 we top-code childcare spending at the 99th percentile value ($1,800 per month), and in column 3-3 we take the median childcare spending instead of the mean. The results are attenuated somewhat when outliers are restricted, with estimates ranging from a reduction of $50 to $56 per month. Scaled up, these coefficients imply an average reduction of roughly $450 to $500 in childcare expenses for the 9-month academic year.[42]

42. The estimate in the final column (3-3) is similar when we limit the comparison group to other southern states (see online appendix table C-2).

While such an effect might seem too small to care much about, recall that this is the reduction in *average* childcare expenses in the population at large, regardless of whether a family takes up universal preschool. Moreover, among those taking up the program, some families—those switching out of private preschools—may see very large reductions in childcare spending, while other families might not be affected much at all. Although we cannot identify the families whose children switch from private to public preschools in the CEX (and would not want to because of endogeneity concerns), we can do a back-of-the-envelope calculation of the magnitude of the transfer to these families based on our private enrollment estimates from the October CPS. Returning to table 2, lower panel, we see declines in private school enrollment of 8 to 14.9 percentage points among 4-year-olds with mothers in the higher education category. Assuming that the effects on childcare expenses in the CEX for this group are explained completely by the movement of 4-year-olds out of private preschool, the top-coded estimates from column 3-2 imply that families that switch from private preschools to the public program save between roughly $3,000 and $5,600 per academic year on childcare expenses.[43] By comparison, the maximum Earned Income Tax Credit in the 2012 tax year for a family with two qualifying children was $5,236. However, universal preschool provides an income transfer only to relatively high-income families.

Indeed, returning to table 3, we see that childcare expenses have not fallen in families where mothers have lower levels of education, which is consistent with the lack of movement out of private preschool for this group in response to high-quality universal preschool. In fact, the gap in childcare expenses between families with 4-year-olds and families with 5-year-olds is higher in Georgia and Oklahoma than it is elsewhere in the country, although this difference disappears when we estimate the differences on median spending (column 3-3).

III.B. Maternal Labor Supply

The enrollment and childcare spending estimates suggest that some families—relatively high-SES families that arguably place a high value on preschool education—receive sizable income transfers from universal preschool programs. However, regardless of income, these programs and public education generally provide a 100 percent price subsidy for childcare on

43. These figures are in line with estimates of childcare expenses from the Survey of Income and Program Participation, which range from about $110 to $143 per week (in real 2011 dollars) over the CEX sample period (Laughlin 2013).

the extensive margin of employment, increasing primary caregivers' (historically, mothers') take-home pay and so in theory providing them with a strong incentive to enter paid work. Increases in maternal employment can increase family income, expanding a family's consumption possibilities as well as its child investment possibilities.

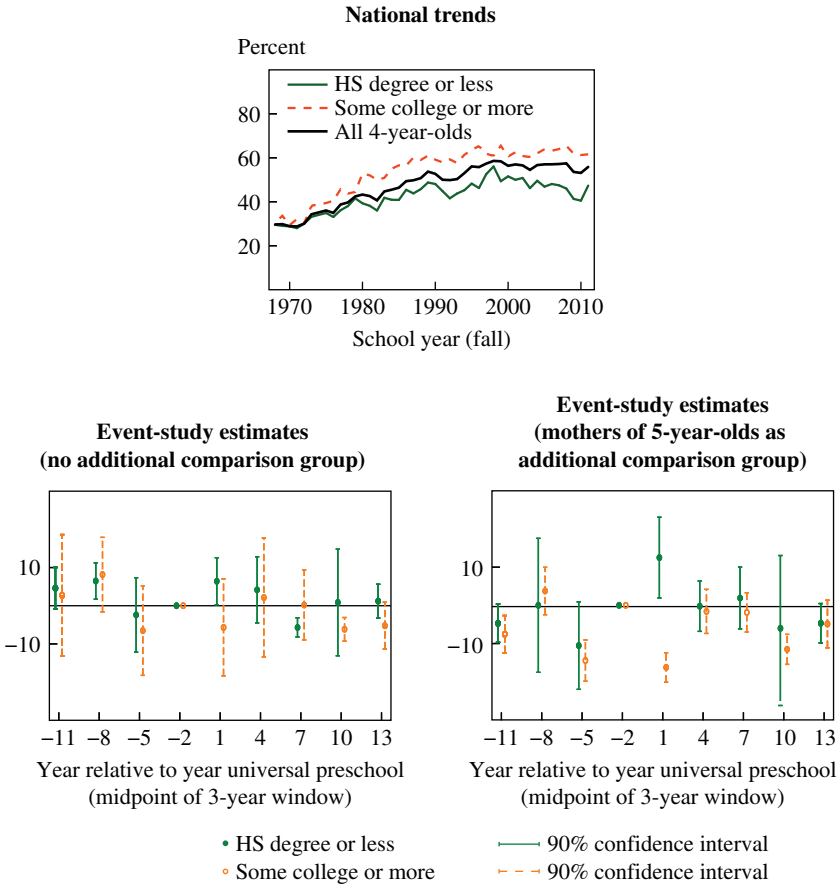
How have the Georgia and Oklahoma programs affected maternal labor supply? To address this question, we return to our match of 4-year-olds to their mothers in the October CPS and to our original DD estimation strategy. Our outcome of interest is the percent of mothers of 4-year-olds who were working in the week prior to the survey.[44] Figure 5 presents national trends in this variable, both overall and by maternal education (top panel), alongside the event-study estimates by maternal education category (lower-left panel). The first two columns of table 4 present the estimates of the $post_{st}$ coefficient from a model analogous to that in equation 1, both without state-by-year varying controls (column 4-1) and with them (column 4-2).

These initial results provide little evidence to suggest that the preschool programs in Georgia and Oklahoma have increased the likelihood that mothers of 4-year-olds are at work. In fact, the coefficient estimates are negative. This is unexpected, and it suggests that a simple DD model like that in equation 1 might not be identifying the effects of the program. Indeed, the event-study estimates in figure 5 (lower-left panel) do not provide compelling evidence of a negative impact. Rather, they suggest that we are finding a negative impact due to unusually high maternal employment rates in Georgia and Oklahoma well before the states introduced universal preschool programs (for example, the coefficient for 7 to 9 years prior to the initiative, plotted at –8, is positive and large). Thus, what appear to be negative program impacts may actually reflect a convergence of the rest of the country to Georgia's and Oklahoma's higher initial rates of maternal employment.

One approach to this problem would be to re-estimate the simple DD models, dropping or dummying up the earliest years of data so that they do not contribute to identification. Another approach, which we find more compelling, is to attempt to account for any such convergence using the mothers of other young children—such as the mothers of 5-year-olds—as an additional comparison group. In theory, these mothers should have experienced the same labor market shocks as the mothers of 4-year-olds, though their own children are not eligible for universal preschool. To the

44. We base our analysis on the employment status recode variable, which has been consistently asked of CPS respondents over our sample frame.

Figure 5. Trends in Employment Rates of Mothers of 4-Year-Olds and Event-Study Estimates of the Effect of Introducing High-Quality Universal Preschool on Maternal Employment, by Maternal Education,[a] 1970–2011



Source: October CPS school enrollment supplements, 1977–2011.

a. States with high-quality preschool programs are Georgia (introduced in fall 1995) and Oklahoma (introduced in fall 1998). Regressions underlying the coefficients in the lower-left panel include state fixed effects and year fixed effects in addition to dummies for 3-year groupings of year relative to the year that high-quality universal preschool was introduced. The coefficients plotted at –11 represent 10 years or more prior to introduction, while the coefficients plotted at 13 represent 12 or more years after; the dummy at –2, representing the three years immediately prior to introduction, is omitted to identify the model. Regressions underlying the coefficients in the lower-right panel include all of the same controls, entered both directly and interacted with an indicator for whether the mother has a 4-year-old. The coefficients plotted are the coefficients on the interactions with the indicators for year relative to the initiative. All regressions were weighted by the number of observations used to calculate the maternal employment rate, and standard errors were clustered on state.

Table 4. Difference-in-Differences and Triple-Difference Estimates of the Impact of High-Quality Universal Preschool on Employment Rates of Mothers of 4-Year-Olds, by Maternal Education[a]

	Coefficient on post (Model 1)		*Coefficient on post × age 4 (Model 3), comparison group is (mothers of)*
Dependent variable	*Baseline 4-1*	*Add controls 4-2*	*5-year-olds 4-3*
	Mother has high school degree or less		
Maternal employment:			
Mothers at work	–1.874	–1.114	4.728**
prior week (%)	(2.533)	(2.700)	(2.273)
Child's school enrollment:			
Public school	19.94***†	18.91***†	15.00***
enrollment rate (%)	(1.408)	(1.386)	(3.091)
Private school	–1.361	–0.638	0.793
enrollment rate (%)	(1.137)	(0.820)	(1.942)
Effect of public on private	*–0.0683*	*–0.0337*	*0.0529*
	(0.0544)	*(0.0424)*	*(0.140)*
N (state-by-year-by-age cells)	1,785	1,785	3,570
	Mother has some college or more		
Maternal employment:			
% Mothers at work prior week	–4.620***	–4.518***	–1.111
	(1.389)	(1.394)	(3.182)
Child's school enrollment:			
Public school	23.32***††	23.35***††	16.61*
enrollment rate (%)	(5.746)	(6.000)	(8.718)
Private school	–15.18***††	–14.90***††	–6.980***
enrollment rate (%)	(1.542)	(0.989)	(1.766)
Effect of public on private	*–0.651****	*–0.638****	*–0.420****
	(0.216)	*(0.189)*	*(0.133)*
N (state-by-year-by-age cells)	1,784	1,784	3,568
Controls:			
State fixed effects	Y	Y	Y
Year fixed effects	Y	Y	Y
Nonwhite (%), female (%)	N	Y	Y
State unemployment rate	N	Y	Y
All controls × *age4* indicator	N	N	Y

Source: Authors' regressions based on state-by-year-by-age averages of microdata from the October Current Population Survey, 1977–2011.

a. See the text for further details on model (1) and model (3). Coefficients in italics are two-stage least squares estimates of the effect of the public school enrollment rate on the private school enrollment rate, where the instrument for the public school enrollment rate is either *post* (in columns 4-1 and 4-2) or *post* × *age4* (in column 4-3). Regressions are weighted by the number of children used to calculate the enrollment rate. Standard errors clustered on state are in parentheses. Asterisks indicate statistical significance at the ***1 percent, **5 percent, and *10 percent level. Daggers indicate statistical significance based on Conley and Taber (2011) confidence intervals at the †† 5 percent or † 10 percent level (and apply to non-italicized DD estimates in columns 4-1 and 4-2 only).

extent that the employment trends of these two groups of mothers diverge after the initiatives were passed in Georgia and Oklahoma, relative to the rest of the United States, it would suggest that the program is having a positive employment impact.

We therefore combine the two empirical approaches that we have used thus far into a "triple-difference" (DDD) model, one that uses variation in the availability of universal preschool across states over time, and across women with children of different ages. The model is given by this equation:

$$y_{ast} = \theta\, post_{st} \times age4_a + \lambda\, post_{st} + \gamma_{as} + \delta_{at} + \varepsilon_{ast}, \tag{3}$$

where y_{ast} is the percent of mothers of children age a in state s who reported working in October of year t; and $post_{st}$ and $age4_a$ are as defined earlier. The vectors γ_{as} and δ_{at} represent age-by-state and age-by-year fixed effects, which control for fixed differences across states in employment rates of mothers with children of different ages and national shocks to employment rates of mothers of children of the same age. The $post_{st}$ direct effect, λ, is the "effect" of universal preschool on women with age-ineligible children. The interaction coefficient, θ, is then the effect of interest, capturing the *differential* impact of universal preschool on the employment of women with 4-year-olds.[45] The identifying assumption is that there are no other reasons to see a difference in employment between women with 4-year-olds and women with age-ineligible children in Georgia and Oklahoma after their universal preschool programs were in place.[46]

The remaining column of table 4 shows estimates of θ in equation 3 using mothers of 5-year-olds as a comparison group. The estimates are quite different than the findings from the basic DD specification. Regardless of maternal education, the estimates are much less negative than those in the first two columns. This suggests that the use of a comparison group helps to account for that convergence suggested in figure 5 (lower-left panel), and indeed, when we examine the triple-difference event-study that uses

45. Cascio (2009b) uses a similar DDD approach to estimate the effect of state funding for kindergartens (and kindergarten enrollment) on the employment of women with 5-year-olds. Most of the kindergarten funding initiatives had passed by the time the data used here began (1977).

46. Note that, if an inter-temporal labor supply framework applies, a mother who works more due to the availability of universal preschool when her child is 4 years old may work less the next year due to a wealth effect, possibly leading us to exaggerate the impact of universal preschool on maternal labor supply in a DDD framework. Because we are looking at the extensive margin of employment (i.e., working versus not) rather than the intensive margin (i.e., hours worked), we think that this source of bias is unlikely.

mothers of 5-year-olds as a comparison group (lower-right panel), the pre-initiative coefficients fluctuate around zero.[47]

Turning to the DDD estimates themselves, we see no impact of universal preschool on the employment rate of women with at least some college education (lower-right panel). The DDD estimate is in fact negative. Importantly, this is not because the DDD specification fails to yield an impact of universal preschool on school enrollment: as shown lower down in the panel, the introduction of universal preschool is associated with a 17-percentage-point increase in 4-year-olds' public school enrollment rates—a figure that, along with the degree of substitution between public and private enrollment, lines up well with the simple DD estimates in table 2.[48] Nevertheless, the estimate is imprecise enough that we cannot rule out positive employment impacts of a reasonable magnitude.[49]

By contrast, the DDD estimate for less-educated mothers is positive and statistically different from zero using conventional methods of inference; for their children, the evidence clearly shows an impact of the programs on school enrollment. While these findings are promising, the corresponding event-study estimates are less compelling. As shown in the lower-right panel of figure 5, the positive coefficient appears to be driven by an impact in the first three years of the program (the solid coefficient plotted above 1). All remaining post-initiative event-study coefficients are statistically insignificant, and some are negative. Further, the DDD employment estimates are sensitive to changes in how we stratify the data, our choices of

47. To arrive at these event-study estimates, we replace $post_{st}$ in model (3) with a full set of indicators for year relative to the initiative (in 3-year bins), and we replace $post_{st} \times age_a$ with interactions between these indicators and age_a. Figure 5, lower-right panel, plots the estimated coefficients on these interaction terms, separately by maternal education.

48. We look at public and private *school* enrollment here, rather than public and private *preschool* enrollment, because most 5-year-olds would be expected to be in kindergarten. From the mother's perspective, however, what matters for maternal employment should be whether a child is in (public) school, not the child's grade of enrollment. To the extent that these numbers deviate in any meaningful way from those in table 2, it suggests an impact of universal preschool programs on the grade in which a child is enrolled. For example, some children may have attended private kindergarten in the absence of public preschool.

49. If we assume that the implicit childcare subsidy from universal preschool is equivalent to an 8 percent net wage increase, a 50 percent baseline maternal employment rate (see table 1), and wage elasticity of labor supply for women ranging from 0.2 to 0.4 (McClelland and Mok 2012), then we would expect that universal preschool would increase the likelihood of working by only 1.6 to 3.2 percentage points.

comparison states and the comparison age group, and weighting.[50] Overall, these findings suggest that the Georgia and Oklahoma programs may have increased maternal employment, but the evidence is not very compelling or robust.

III.C. Maternal Time with Children

If preschool is a substitute for childcare, then the time parents spend on childcare should decline as the time their children spend in preschool increases. Accordingly, given little evidence that lower-SES families switch out of private programs, we would expect relatively large reductions in the time mothers spend with children in these families. However, universal programs might also change optimal time investments in children. For example, preschool exposure might encourage parents to spend more or less time engaging in educational activities with their children, such as reading to them and doing art projects.

To investigate this question, we use data from the American Time Use Surveys spanning January 2003 through December 2012. These surveys provide nationally representative estimates of how, where, and with whom respondents spend their time. Because they contain no pre-initiative data, we take the same general approach as in the CEX, comparing differences across mothers with 4-year-olds and mothers with 5-year-olds on daily time spent on childcare, in the two model states versus the rest of the country, using equation 2.[51] Note that as with the CEX, we cannot isolate time spent on an individual child when there are other children in the household. Thus, these data share many of the limitations of the CEX, so our conclusions from them are, at best, suggestive.

In table 5, column 5-1 shows the impact on the overall amount of time a mother spends in the presence of her child.[52] Regardless of education, mothers spend an average of almost 8 hours per day with their children. In the absence of universal preschool, mothers also spend more time with

50. See online appendix tables B-4 and B-6 (for results that stratify the data on eligibility for free or reduced-price lunch and on paternal college attendance, respectively); online appendix table C-3 (for results that include only southern states in the comparison group); online appendix table G-2 (for results that use mothers of 3-year-olds as a comparison group); and online appendix table D-3 (for unweighted estimates).

51. We limit the data set to women co-residing with a 4- or 5-year-old child, and exclude time-use observations collected on weekends, holidays, or during July or August. To increase statistical power, we average within state-by-education cell across all years pooled together.

52. Activities for which the American Time Use Survey does not collect information on who else was present, such as sleeping, are omitted from this measure.

Table 5. Difference-in-Differences Estimates of the Impact of High-Quality Universal Preschool on Mothers' Time Use, by Maternal Education[a]

Coefficient on:	*Time spent with child present* 5-1	*Time spent caring for and/or helping child* 5-2
	a. Mother has high school degree or less (N = *96)*	
Age 4 × GA or OK	–46.18*	24.96**
	(24.65)	(11.59)
Age 4	32.74***	16.54**
	(11.80)	(6.693)
GA or OK	27.95	–15.65
	(41.25)	(10.95)
Constant	477.3***	101.6***
	(12.12)	(5.787)
	b. Mother has some college or more (N = *102)*	
Age 4 × GA or OK	2.702	–7.703
	(44.41)	(7.352)
Age 4	13.29	15.51***
	(10.50)	(4.867)
GA or OK	–1.274	–7.849
	(34.52)	(6.122)
Constant	461.8***	121.5***
	(7.781)	(3.408)

Source: Authors' regressions based on state-by-presence of 4-year-old averages of microdata from the American Time Use Survey, January 2003–December 2012.

a. Each panel and column represents a separate estimate of model (2). Time is measured in minutes, and the sample is limited to women in households with a 4- or 5-year-old child, and to non-holiday weekdays from September through June. Regressions are weighted by the number of families used to calculate the dependent variable. Standard errors clustered on state are in parentheses. Asterisks indicate statistical significance at the ***1 percent, **5 percent, or *10 percent level.

4-year-olds than with 5-year-olds, though this difference is only statistically significant for less-educated moms (upper panel). With universal preschool, however, this gap is eliminated: the DD coefficient indicates that less-educated mothers in Georgia and Oklahoma on average spend 46 fewer minutes per weekday in the presence of their 4-year-olds than their counterparts elsewhere in the country—an effect amounting to around 4 hours a day when rescaled by the impacts of universal preschool on overall school enrollment.[53] There is not a significant reduction in the time that

53. This is more than would be expected if the preschool programs were only half-day. However, in auxiliary regressions, we find that much of the impact of the Georgia and Oklahoma programs on preschool enrollment, regardless of maternal education, can be accounted for by full-day programs.

more-educated mothers spend in the presence of 4-year-olds, but the DD estimate is noisy (lower panel).

In column 5-2, we turn to the impact on time spent caring for and helping the household's children, which includes time spent with children reading, playing, doing art projects, and talking. The program increases maternal time spent in these activities by 25 minutes per day, which is consistent with experimental evidence from Head Start (Gelber and Isen 2013).[54] Among women with high levels of education, time spent caring for and helping children does not appear to be affected.

IV. Standardized Test Scores

The human capital benefits of high-quality universal preschool are likely to accrue largely from the preschool experience relative to the alternative use of a child's time. Because of the findings for enrollment and crowd-out, we expect to see clear benefits for lower-SES children: the marginal public preschool enrollee would otherwise not have attended preschool, the programs themselves meet high-quality benchmarks, and exposure to higher-income peers in a universal preschool classroom may have additional positive impacts. We also find suggestive evidence that parental investments in children rise in lower-income families. For higher-SES children, the potential effects are less clear: the marginal attendee has a high probability of attending a private preschool, that private preschool may be of equal or better quality than the state-funded one, and the attendee's academic achievement might be diminished by negative peer effects in the classroom. Higher-SES children may also gain less in academic performance if there are diminishing returns to human capital investment.

To measure whether test scores improve in response to the Georgia and Oklahoma programs, we assembled state-by-year aggregate data from the NAEP, which is the only standardized test that is comparable across the nation. Students are tested in mathematics and reading in grades 4 and 8 in selected years. Information is not available on maternal educational attainment in these students' families, so we separate the sample instead by whether a student's family earns less than 185 percent of the poverty line

54. This finding continues to hold when we limit the comparison group to southern states (online appendix table C-4) and when we give each state-by-age observation equal weight (online appendix table D-4). However, it is no longer statistically different from zero when we use confidence intervals that take into account the fact that we have only two treatment states (online appendix table E-4).

and is therefore eligible for free or reduced-price lunch. As shown in online appendix tables B-3 and B-4, the overall pattern of findings for enrollment and maternal labor supply is substantively similar to that shown in tables 2 and 4 when we stratify the data by a child's free or reduced-price lunch eligibility.

We focus on state-by-year mean scale scores on the NAEP math and reading tests. Figure 6 presents trends in mean national scale scores for math and reading, plotting fourth and eighth grade scores on the same graph against the preschool cohort to which they pertain. For example, the earliest available data at the national level are from spring 1996; these data correspond to children who would have been of preschool age in fall 1990 in the case of fourth grade scores and in fall 1986 in the case of eighth grade scores.[55] The preschool cohort is given on the horizontal axis, and the year of the test is given above each data point. Unsurprisingly, eighth graders score higher on this vertically scaled test, and lower-income students perform relatively poorly in each grade. Test scores have improved over time, albeit with little reduction in the test score gap by family income, with the most notable improvements coming in fourth-grade scores for the cohorts that turned age four around the time the Georgia and Oklahoma programs went into effect.

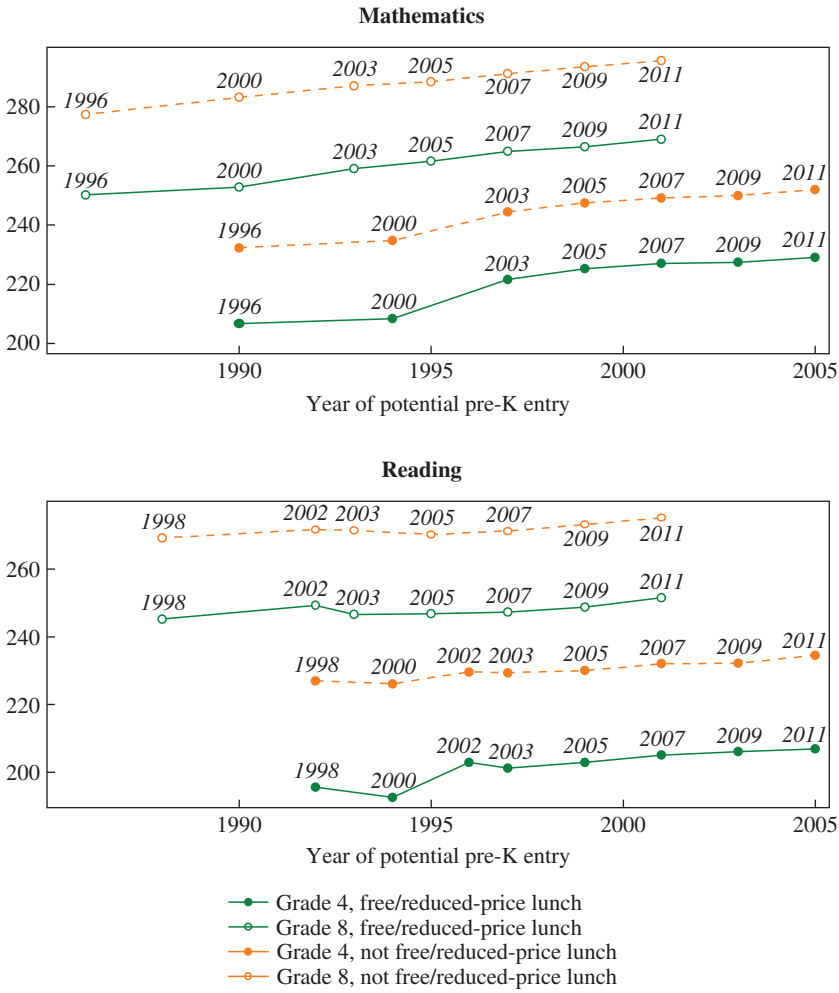
Recall that our preferred empirical approach, used to estimate impacts on preschool enrollment and maternal employment, has been to compare trends in Georgia and Oklahoma relative to the rest of the country, before and after the introduction of their initiatives, and that simple DD estimates rely on the assumption that these trends are similar in the treatment and comparison states in the pre-initiative period. Figure 7, which plots event-study estimates, provides evidence that this assumption is violated for NAEP scores.[56] For example, fourth grade math scores are significantly lower than would be expected given trends elsewhere in the country in the years immediately before the initiative (upper-left panel); fourth grade reading scores show a similar pattern, though the differences across pre-initiative years are not statistically significant (upper-right panel).[57] On the other

55. This calculation is made under the assumption of normal grade progression. Data are available by state and family income starting in spring 2000 (math) and spring 1998 (reading).

56. Here, we use two-year bins. As above, however, the first and last indicators correspond to all prior and subsequent years and we omit the identifier for the year(s) immediately prior to the initiative to identify the model.

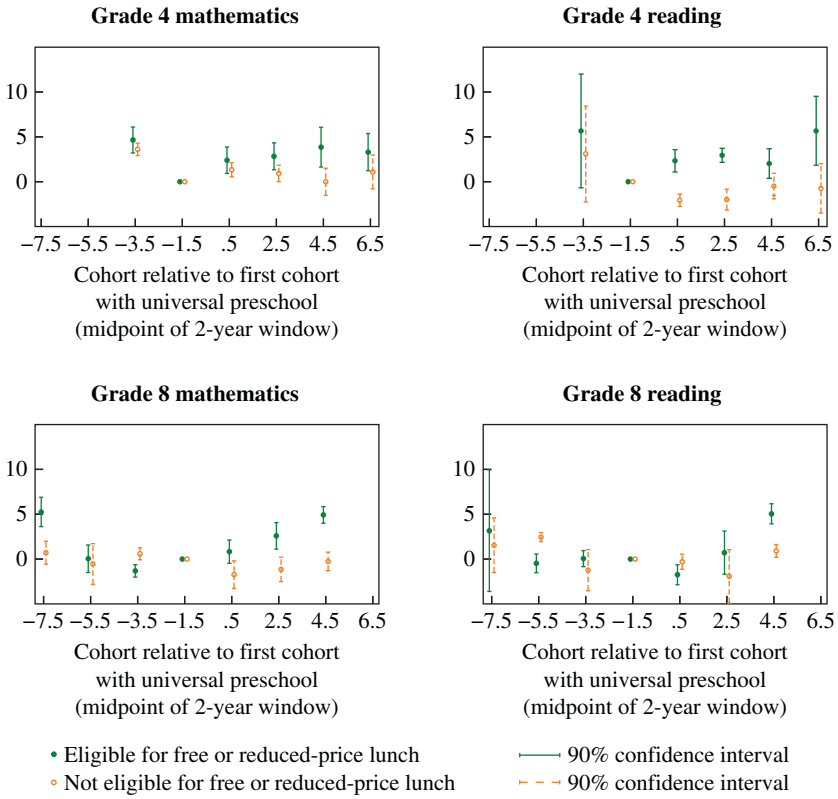
57. Fitzpatrick (2008) found little evidence that fourth grade NAEP scores in Georgia were on a different trajectory from the rest of the nation in the years before introducing universal preschool. Consistent with this, auxiliary analysis suggests that Oklahoma is the source of pre-initiative trends in figure 7.

Figure 6. National Trends in Mean NAEP Scale Scores, 1996–98 to 2011, by Grade, Subject, and Eligibility for Free or Reduced-Price Lunch[a]



Source: Main NAEP Data Explorer (NDE).
a. Test years (spring) are given above the data points.

Figure 7. Event-Study Estimates of the Effect of Introducing High-Quality Universal Preschool on Fourth and Eighth Grade State Mean NAEP Scores, by Eligibility for Free or Reduced-Price Lunch[a]



Source: Main NAEP Data Explorer (NDE). State mean scores by free/reduced-price lunch status in math are available for spring 2000, 2003, 2005, 2007, 2009, and 2011. Mean scores by free/reduced-price lunch status in reading are available for spring 1998, 2002, 2003, 2005, 2007, 2009, and 2011. Grade 4 (8) scores pertain to cohorts of age to enter preschool the fall 6 (10) years prior, assuming normal grade progression.

a. States with high-quality preschool programs are Georgia (introduced in fall 1995) and Oklahoma (introduced in fall 1998). All regressions include state fixed effects and cohort fixed effects in addition to dummies for 2-year groupings of cohorts relative to the first cohort exposed to high-quality universal preschool. The first and last coefficients in each graph are on dummies set to one for all prior and later cohorts, respectively; the dummy at –1.5, representing the two cohorts immediately prior to introduction, is omitted to identify the model. Regression standard errors were clustered on state.

hand, when we look at eighth grade test scores (lower panels), the evidence suggests that the differences in test performance between Georgia and Oklahoma and the rest of the country were largely confined to cohorts that were of preschool age well before these state programs were being introduced. For example, lower-income children in Georgia and Oklahoma who were 4 years old seven or more years prior to the introduction of universal preschool scored higher in eighth grade math and reading than their counterparts elsewhere in the country. However, this advantage no longer existed among later cohorts that were 4 years old prior to the initiative, and among cohorts eligible for universal preschool the advantage to be being in Georgia or Oklahoma returned, suggesting that the program had an impact.

Table 6 presents DD estimates for fourth grade NAEP scores, separately for lower- and higher-income children. None of basic DD estimates is statistically significant, and the estimates for higher-income children are in fact negative (column 6-1), which is consistent with the graphical evidence of strong downward trends in test scores in the treated states prior to their establishing universal preschool. However, when we attempt to account for the early pre-initiative differences in test scores by adding a dummy for the cohorts 3 or more years prior to the initiative (column 6-2)—and thus identify the DD coefficient off of test scores immediately before the initiatives were passed—the estimates become more positive, and for lower-income students they are marginally statistically different from zero for math and reading scores alike (upper panel). An alternative approach, which controls for state-specific trends (column 6-3), moves the coefficients in the same direction but to a lesser extent. This model does not seem as good a characterization of the patterns shown in figure 7, however, so we return to the specification in column 6-2 in the remaining columns of the table, adding controls for student demographics (column 6-4) and years that a state has had "consequential" school accountability (column 6-5).[58] Neither of these sets of controls appreciably changes the point estimates relative to those in column 6-2.

Table 7 shows a similar pattern of findings for eighth grade NAEP scores, although the point estimates are smaller in magnitude. Based on the fully

58. The school accountability programs put in place after the No Child Left Behind (NCLB) Act of 2001, as well as the "consequential" state accountability programs that preceded NCLB, may be an alternative explanation for the patterns we see in the data. Oklahoma implemented consequential accountability in 1996 and Georgia in 2000 (Dee and Jacob 2011). Comparing states that implemented consequential accountability as a result of NCLB to those that had it prior to NCLB, Jacob and Dee (2010) and Dee and Jacob (2011) find evidence that NCLB resulted in higher NAEP scores, particularly in fourth grade.

Table 6. Difference-in-Differences Estimates of the Impact of High-Quality Universal Preschool on Fourth Grade NAEP Mathematics and Reading Scores, by Child's Eligibility for Free or Reduced-Price School Lunch[a]

	Effect sizes				
Dependent variable	*Baseline DD* 6-1	*Baseline + indicator for 3+ cohorts pre-initiative* 6-2	*Baseline + add state linear trends* 6-3	*Specification 2 + student demographics* 6-4	*Specification 4 + other state education policies* 6-5
	Child eligible for free or reduced-price lunch				
Math scale score, 4th grade	1.769	3.201***	2.643*	3.146**	3.138**
	(1.791)	(1.189)	(1.332)	(1.182)	(1.214)
N (state-years)	296	296	296	296	296
Reading scale score, 4th grade	0.398	3.288***	2.468	3.011***	3.154***
	(2.118)	(1.095)	(2.472)	(1.083)	(1.155)
N (state-years)	339	339	339	339	339
	Child not eligible for free or reduced-price lunch				
Math scale score, 4th grade	–0.256	0.834	0.467	0.725	0.857
	(1.130)	(0.627)	(1.089)	(0.593)	(0.719)
N (state-years)	296	296	296	296	296
Reading scale score, 4th grade	–2.968	–1.332	–0.653	–1.585*	–1.589*
	(1.939)	(0.879)	(0.864)	(0.853)	(0.848)
N (state-years)	339	339	339	339	339
Controls:					
State fixed effects	Y	Y	Y	Y	Y
Cohort fixed effects	Y	Y	Y	Y	Y
3+ cohorts prior to first affected cohort (=1)	N	Y	N	Y	Y
Linear trends for GA and OK	N	N	Y	N	N
Student demographics	N	N	N	Y	Y
Years with school accountability	N	N	N	N	Y

Source: Authors' regressions based on state-by-year test score data from NAEP Data Explorer, data on enrollment by race from the Common Core of Data, and dates of consequential school accountability reported in Dee and Jacob (2011).

a. Each reported coefficient is a separate estimate of θ from model (1). See the text for details. Standard errors clustered on state are in parentheses. Asterisks indicate statistical significance at the ***1 percent, **5 percent, or *10 percent level.

Table 7. Difference-in-Differences Estimates of the Impact of High-Quality Universal Preschool on Eighth Grade NAEP Mathematics and Reading Scores, by Child's Eligibility for Free or Reduced-Price School Lunch[a]

	Effect sizes				
Dependent variable	*Baseline DD* *7-1*	*Baseline + indicator for 7+ cohorts pre-initiative* *7-2*	*Baseline + add state linear trends* *7-3*	*Specification 2 + student demographics* *7-4*	*Specification 4 + other state education policies* *7-5*
	Child eligible for free or reduced-price lunch				
Math scale score, 8th grade	2.015	2.724*	2.455**	2.252*	2.152*
	(1.938)	(1.434)	(1.215)	(1.197)	(1.121)
N (state-years)	295	295	295	295	295
Reading scale score, 8th grade	0.193	1.014	0.155	0.899	0.819
	(2.729)	(1.912)	(1.393)	(1.873)	(1.874)
N (state-years)	334	334	334	334	334
	Child not eligible for free or reduced-price lunch				
Math scale score, 8th grade	–1.115	–1.009	–1.756***	–1.260	–1.293
	(1.033)	(0.983)	(0.360)	(0.850)	(0.835)
N (state-years)	295	295	295	295	295
Reading scale score, 8th grade	–1.078	–0.763	0.111	–0.769	–0.813
	(2.387)	(2.304)	(1.061)	(2.294)	(2.297)
N (state-years)	334	334	334	334	334
Controls:					
State fixed effects	Y	Y	Y	Y	Y
Cohort fixed effects	Y	Y	Y	Y	Y
7+ cohorts prior to first affected cohort (=1)	N	Y	N	Y	Y
Linear trends for GA and OK	N	N	Y	N	N
Student demographics	N	N	N	Y	Y
Years with school accountability	N	N	N	N	Y

Source: Authors' regressions based on state-by-year test score data from NAEP Data Explorer, data on enrollment by race from the Common Core of Data, and dates of consequential school accountability reported in Dee and Jacob (2011).

a. Each reported coefficient is a separate estimate of θ from model (1). See the text for details. Standard errors clustered on state are in parentheses. Asterisks indicate statistical significance at the ***1 percent, **5 percent, or *10 percent level.

controlled specification, in column 7-5, the effect of high-quality universal preschool on math scores declines from 3.1 points in fourth grade to (a marginally statistically significant) 2.2 points in eighth grade. For reading, the decline is even more extreme—from 3.1 points to 0.8 points between fourth and eighth grade. This pattern of findings is consistent with the pervasive finding of test score "fadeout" in the broader preschool evaluation literature, discussed in section I.C. When viewed in this context, the persistence of test score impacts through eighth grade is somewhat surprising.

The point estimates from table 7 indicate that lower-income children who were eligible for universal preschool in Georgia and Oklahoma scored on average about 2.2 points higher on the NAEP math test in eighth grade than would have been expected based on national trends, after accounting for several other state-by-year varying factors. Depending on the estimate, this amounts to a gain of 0.058 to 0.067 student-level standard deviation. Rescaled by the programs' impacts on the preschool enrollment rates of lower-SES children and assuming no positive spillover effects onto classmates, these findings suggest that preschool attendance increases eighth-grade math scores by almost a third of a standard deviation. For fourth grade math and reading alike, the estimates are even larger.[59] Given our findings for maternal time use and labor supply, we think it unlikely that the impacts are working through preschool attendance alone.

That said, as with our findings for maternal employment, our findings for NAEP scores are not as robust to alternative specifications as those for enrollment. For example, the DD estimates are substantially smaller when we limit the comparison group to states in the South (see online appendix tables C-5 and C-6).[60] They are also not statistically significant at conventional levels when we account for the fact that we have only two treatment states (online appendix tables E-5 and E-6). Based on the Conley and Taber (2011) methodology, the most precise test score results are for fourth grade math; these are statistically different from zero at the 20 percent level.

59. The fourth grade estimates are thus larger than those found by Fitzpatrick (2008) in her study of Georgia. One possible explanation is that we have three more years of post-initiative data at our disposal (from 2007, 2009, and 2011), and the reduced-form program impacts appear to grow slightly over time (figure 7, upper panels). However, the studies are not necessarily comparable since we include Oklahoma as a treatment state. Unfortunately, we lack sufficient pre-initiative observations in the state-aggregated NAEP (that is, we lack 1996 data by free-lunch status) to present informative estimates separately by state.

60. When rescaled by the slightly smaller enrollment impacts when the comparison group is limited to the South (see online appendix table C-1), the implied effect of public preschool enrollment at age 4 on later test scores is closer in magnitude.

V. Cost-Benefit Analysis

Economists typically attempt to quantify potential benefits of social programs in order to compare them to costs and judge whether they make a worthwhile investment. There is necessarily a lot of uncertainty in projecting the benefits of universal, high-quality preschool into the future; the projected benefits depend on a variety of assumptions about the state of the labor market in the future and about the discount rate. Our estimates of benefits, like our estimates of test score impacts, are also less precise than would be ideal for an exercise such as this. Below we attempt to quantify the projected long-term impacts of preschool, recognizing these limitations.

V.A. Measuring Benefits

To predict the long-term earnings impact of high-quality universal preschool, we predict how improved math test scores will increase future earnings over the life span. The first step in this prediction is to estimate the impact on the math scores of all children, which we find to be 2.4 points in fourth grade and (a statistically insignificant) 0.9 points in eighth grade. The next step is to convert these "intention-to-treat" estimates, which combine outcomes for children whether or not they attended preschool, into "treatment-on-treated" impacts, which rescale by the program's impacts on public preschool enrollment—a 23.27-percentage-point increase in the pooled sample.[61] Using this scaling factor, we arrive at treatment-on-treated estimates equal to 0.29 and 0.11 standard deviations in grades 4 and 8, respectively.

Next, we convert this test score improvement into predicted wage gains using the relationship between earnings and test scores. The best available estimate of this relationship is from Raj Chetty, John Friedman, and Jonah Rockoff (2013), who find that a 1 standard deviation improvement in test scores (measured in grades 4 through 8) correlates to a 12 percent increase in earnings when a worker is in his or her mid- to late-20s. Assuming that this relationship continues to hold across the rest of the worker's career, this estimate implies annual earnings impacts of 3.5 percent (when benefits are based on the fourth grade math results) and 1.3 percent (when benefits are based on the eighth grade math results). To convert this wage boost into dollar terms across a worker's career, we apply it to the age-earnings

61. One drawback to this approach is that it assumes there are no positive spillover impacts of preschool attendance onto a child's classmates who did not attend preschool.

profile of workers with positive earnings using the 2011 March CPS. Following the Congressional Budget Office's long-term forecast, we assume real labor productivity growth equal to 1.9 percent per year. We also assume an 80 percent labor force participation rate and that workers are employed from age 18 to age 65.

V.B. Measuring Costs

Estimating the costs of preschool is arguably more straightforward. The most recent available data report annual costs per enrollee of $3,652 in Georgia and $7,427 in Oklahoma. Since the impacts are estimated based on the pooled impact across these two states, we take the enrollment-weighted average of per-pupil costs as a total cost measure. This weighted average is $4,698 and represents the total government outlays for the program. Importantly, though, this overstates the true economic costs, because some of this spending offsets the out-of-pocket costs that would have been borne by families who would have sent their children to preschool in the absence of the public program. In other words, some of this spending is not new spending, but can be thought of as an income transfer to parents who would otherwise pay for preschool out of pocket. Our most conservative estimate of the share of preschoolers who substitute from private to public preschool is 16 percent. Once this income transfer is subtracted, the estimated spending per student is $3,946.

V.C. Comparing Present Discounted Values

Table 8 gives present discounted values of lifetime earnings using a range of discount rates and compares these to costs. The upper panel projects benefits from the point estimate for fourth grade test scores, and the lower panel projects them from the (statistically insignificant) eighth grade point estimate. In column 8-1 we discount using the return on the 30-year Treasury bill, which is the government's long-term borrowing rate. The 30-year interest rate has averaged 3.4 percent from January through October 2013. At this rate, the present discounted benefits, as projected from the fourth grade test score impacts, equal $33,740. These benefits outweigh government outlays by a factor of seven. When benefits are compared to economic costs less the transfer from the government to families that switch enrollment from private to public preschools, the benefit-cost ratio increases to approximately 8.5. Substituting a higher discount rate reduces the benefits, but even with a 6 percent discount rate, the benefits outweigh the economic costs by 3.4 to 1. As shown in the lower panel, the estimated benefits are more modest based on the eighth grade test score impacts, but

Table 8. Cost-Benefit Analysis of Universal Preschool Program, Based on Test Score Impacts in Fourth and Eighth Grades

	Annual discount rate			
Dependent variable	*3.4 percent 8-1*	*4 percent 8-2*	*5 percent 8-3*	*6 percent 8-4*
	Benefits estimated from 4th grade test score impacts			
Present discounted value	$33,740	$26,997	$18,876	$13,419
Ratio of PDV to total outlays	7.18	5.75	4.02	2.86
Ratio of PDV to net outlays	8.55	6.84	4.78	3.40
	Benefits estimated from 8th grade test score impacts			
Present discounted value	$12,798	$10,240	$7,160	$5,090
Ratio of PDV to total outlays	2.72	2.18	1.52	1.08
Ratio of PDV to net outlays	3.24	2.60	1.81	1.29

Notes: Wage impacts are estimated based on the 2011 age-earnings profile in the Current Population Survey, assuming a 1.9 percent per year real productivity growth rate and an 80 percent labor force participation rate. A 1-standard-deviation increase in math test scores is assumed to increase earnings by 12 percent. Estimated math score impacts are 0.29 and 0.11 standard deviations in grades 4 and 8, respectively. Present discounted values are compared to current-year total outlays of $4,698 and net outlays of $3,946.

when they are discounted by the 30-year Treasury rate they still outweigh net outlays by a factor of 3.2 to 1.

Due to lack of precision and effect fadeout, there is considerable uncertainty surrounding the magnitude of the test score impacts. Another way to think about the cost-effectiveness of the program is to solve for the level of the test score impact that equates the long-term benefits to the costs. Assuming a 3.4 percent annual discount rate, a treatment-on-treated test score impact of 0.04 standard deviations equates the benefits to the total outlays for the program. An impact of 0.03 standard deviations is enough in present discounted terms to offset the net economic costs of the program. These translate to between 1.0 and 1.4 scale score points on the NAEP.

V.D. Additional Considerations

The likely benefits of the program are understated in our analysis, for at least two reasons. First, our benefits calculations are based entirely on projected earnings increases. Other work on early childhood education has found that there are strong long-term impacts on other outcomes, such as criminal behavior and use of the social safety net. In their extraordinarily thorough cost-benefit analysis of the Perry Preschool Program, Heckman and others (2010) find that when benefits that accrue to society more broadly are included, the internal rate of return increases by at least

20 percent over the return to individuals. We do not estimate potential benefits along these other dimensions in our cost-benefit calculation. Second, we predict future earnings increases from fourth and eighth grade test score improvements. Other research on early childhood education has found that realized earnings impacts are larger than those implied by subsequent test score gains alone (Chetty and others 2011). This could be due to improvements in non-cognitive skills, which are rewarded in the labor market but are not well measured by cognitive achievement tests.

VI. Conclusions

President Obama's $75 billion Preschool for All initiative calls for dramatic increases in the number of 4-year-olds in public preschool programs and in the quality of these programs across the nation. His proposal shares—and other proposals likely to follow will share—many characteristics with the state-funded preschool programs in Georgia and Oklahoma, which both meet high-quality benchmarks and are accessible to all children. We estimate the impacts of these model programs on a variety of child and family outcomes using difference-in-differences frameworks. Our findings suggest sharply different impacts for children across the income distribution, which is not surprising when one recognizes that the impact of attending a high-quality public preschool depends crucially on what the child would have been doing in the absence of the program.

For lower-SES children, we find evidence of increases in math scores that may be sustained through eighth grade. The increases may be working through multiple channels. First, children are likely to switch from not attending preschool to attending a high-quality public preschool when a universal program is introduced. Moreover, we find suggestive evidence that although they spend less time overall in the presence of their mothers, these children spend more time actively engaging in activities such as playing and reading with them. We also find suggestive evidence that their mothers are more likely to become employed.

For higher-SES children, we find no positive impacts on student achievement. These children are much less likely to be moved on the extensive margin of preschool enrollment, and instead in response to the program are more likely to switch from private to public preschool. We also find suggestive evidence that some families are able to substantially reduce their spending on private preschool and childcare in response to the program, freeing these resources up for other purposes.

This pattern of results raises the question of whether the proposal design could be altered to obtain the same positive impacts without inducing as much crowd-out. Could a targeted program meet these goals more efficiently? The findings of rapid and complete fadeout in test score effects in recent randomized controlled trials of Head Start (Puma and others 2012) and Tennessee's targeted voluntary preschool program (Lipsey and others 2013b) suggest that targeted programs today might not induce such gains. One possible explanation is that the test score impacts of universal programs rely on peer effects in preschool classrooms. Indeed, universal programs might be "high quality" not because they meet specific quality benchmarks, but rather because of improvements in the classroom environment from the presence of higher-SES children. We cannot rule out this possibility, and we think it is an important question for future research.

That said, while the Georgia and Oklahoma programs share a number of features with the preschool programs proposed under the Obama plan, there are several reasons to be cautious in generalizing our findings. First, and most importantly, the Obama plan gives states discretion over whether and how much to charge middle-class families for access to the public programs. The more a state charges middle-class families, the less substitution away from private programs there is likely to be. Second, although universal preschool has been the focus of our analysis, it is not all that there is to the Preschool for All initiative. For example, the initiative also calls for increases in Head Start enrollment among 3-year-olds, which might affect the success of universal preschool in ways that we have not been able to measure. On the other hand, our suggestion that the programs in Georgia and Oklahoma have had some lasting impacts on the human capital of lower-SES children might have nothing to do with the quality benchmarks emphasized in the Obama plan. We have estimated the reduced-form impacts of these programs, and the mechanisms at work are not completely clear. The available data have also limited our analysis in important ways that might be rectified in future research.

ACKNOWLEDGMENTS We thank Patricia Anderson and Jesse Rothstein; our discussants, Caroline Hoxby and Alan Krueger; the editors; and other participants in the Fall 2013 Brookings Panel on Economic Activity for helpful comments; as well as Mary Zaki, Chase Eck, and Gardiner Kreglow for helpful research assistance.

References

Anderson, Michael L. 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, no. 484: 1481–95.

Barnett, W. Steven, Megan E. Carolan, Jen Fitzgerald, and James H. Squires. 2012. *The State of Preschool 2012: State Preschool Yearbook.* New Brunswick, N.J.: National Institute for Early Education Research.

Bassok, Daphna. 2012. "Raising Teacher Education Levels in Head Start: Are There Program-Level Tradeoffs?" Working Paper no. 3, Center on Education Policy and Workforce Competitiveness, University of Virginia.

———, Maria Fitzpatrick, and Susanna Loeb. 2012. "Does State Preschool Crowd-Out Private Provision? The Impact of Universal Preschool on the Childcare Sector." NBER Working Paper no. 18605. Cambridge, Mass.: National Bureau of Economic Research.

Cascio, Elizabeth. 2009a. "Do Investments in Universal Early Education Pay Off? Long-Term Effects of Introducing Kindergartens into Public Schools." NBER Working Paper no. 14951. Cambridge, Mass.: National Bureau of Economic Research.

———. 2009b. "Maternal Labor Supply and the Introduction of Kindergartens into American Public Schools." *Journal of Human Resources* 44, no. 1: 140–70.

Chetty, Raj, John Friedman, Nathaniel Hilger, Emmanuel Saez, Diane Whitmore Schanzenbach, and Danny Yagan. 2011. "How Does Your Kindergarten Classroom Affect Your Earnings? Evidence from Project STAR." *Quarterly Journal of Economics* 126, no. 4: 1593–1660.

Chetty, Raj, John Friedman, and Jonah Rockoff. 2013. "Measuring the Impact of Teachers II: Teacher Value-Added and Student Outcomes in Adulthood." NBER Working Paper no. 19424. Cambridge, Mass.: National Bureau of Economic Research.

Conley, Timothy and Christopher Taber, 2011. "Inference with 'Difference in Differences' with a Small Number of Policy Changes." *Review of Economics and Statistics* 93, no. 1: 113–25.

Currie, Janet. 2001. "Early Childhood Education Programs." *Journal of Economic Perspectives* 15, no. 2 (Spring): 213–38.

———, and Duncan Thomas. 1995. "Does Head Start Make a Difference?" *American Economic Review* 85, no. 3 (June): 341–64.

Dee, Thomas S., and Brian A. Jacob. 2011. "The Impact of No Child Left Behind on Student Achievement." *Journal of Policy Analysis and Management* 30, no. 3: 418–46.

Deming, David. 2009. "Early Childhood Intervention and Life-Cycle Skill Development: Evidence from Head Start." *American Economic Journal: Applied Economics* 1, no. 3: 111–34.

———, and Susan Dynarski. 2008. "The Lengthening of Childhood." *Journal of Economic Perspectives* 22, no. 3: 71–92.

Espinosa, Linda M. 2002. "High-Quality Preschool: Why We Need It and What It Looks Like." *NIEER Preschool Policy Matters* no. 1 (November). Available at http://nieer.org/resources/policybriefs/1.pdf.

Figlio, David, and Jeffrey Roth. 2009. "The Behavioral Consequences of Pre-Kindergarten Participation for Disadvantaged Youth." In *The Problems of Disadvantaged Youth: An Economic Perspective,* edited by Jonathan Gruber. Chicago and London: University of Chicago Press.

Fitzpatrick, Maria D. 2008. "Starting School at Four: The Effect of Universal Pre-Kindergarten on Children's Academic Achievement." *The B.E. Journal of Economic Analysis & Policy* 8, no. 1.

———. 2010. "Preschoolers Enrolled and Mothers at Work? The Effects of Universal Pre-Kindergarten." *Journal of Labor Economics* 28, no. 1: 51–85.

———. 2012. "Revising Our Thinking about the Relationship between Maternal Labor Supply and Preschool." *Journal of Human Resources* 47: 583–612.

Garces, Eliana, Duncan Thomas, and Janet Currie. 2002. "Longer-Term Effects of Head Start." *American Economic Review* 92, no. 4: 999–1012.

Gelbach, Jonah B. 2002. "Public Schooling for Young Children and Maternal Labor Supply." *American Economic Review* 92, no. 1: 307–22.

Gelber, Alexander, and Adam Isen. 2013. "Children's Schooling and Parents' Behavior: Evidence from the Head Start Impact Study." *Journal of Public Economics* 101: 25–38.

Goldin, Claudia, Lawrence Katz, and Ilyana Kuziemko. 2006. "The Homecoming of American College Women: The Reversal of the College Gender Gap." *Journal of Economic Perspectives* 20, no. 4: 133–56.

Gormley, William T., and Ted Gayer. 2005. "Promoting School Readiness in Oklahoma: An Evaluation of Tulsa's Pre-K Program." *Journal of Human Resources* 40, no. 3: 533–58.

Gormley, William T., Deborah Phillips, and Ted Gayer. 2008. "Preschool Programs Can Boost School Readiness." *Science* 320, no. 5884: 1723–24.

Gramlich, Edward M. 1986. "Evaluation of Education Projects: The Case of the Perry Preschool Program." *Economics of Education Review* 5, no. 1: 17–24.

Heckman, James J. 2006. "Skill Formation and the Economics of Investing in Disadvantaged Children." *Science* 312, no. 5782: 1900–02.

———, Seong Hyeok Moon, Rodrigo Pinto, Peter Savelyev, and Adam Yavitz. 2010. "The Rate of Return to the High Scope Perry Preschool Program." *Journal of Public Economics* 94, nos. 1-2: 114–28.

Heckman, James J., Rodrigo Pinto, and Peter Savelyev. 2013. "Understanding the Mechanisms through Which an Influential Early Childhood Program Boosted Adult Outcomes." *American Economic Review* 103, no. 6: 2052–86.

Jacob, Brian A., and Thomas S. Dee. 2010. "The Impact of No Child Left Behind on Students, Teachers, and Schools." *Brookings Papers on Economic Activity* (Fall): 149–94.

Knudsen, Eric I., James J. Heckman, Judy L. Cameron, and Jack P. Shonkoff. 2006. "Economic, Neurobiological, and Behavioral Perspectives on Building America's Future Workforce." *Proceedings of the National Academy of Sciences* 103, no. 27: 10155–62.

Ladd, Helen F., Clara G. Muschkin, and Kenneth Dodge. 2012. "From Birth to School: Early Childhood Initiatives and Third Grade Outcomes in North Carolina." Working Paper, Duke University.

Laughlin, Lynda. 2013. "Who's Minding the Kids? Childcare Arrangements: Spring 2011." In *Current Population Reports,* P70-135. Washington, D.C.: U.S. Census Bureau.

Lipsey, Mark W., Kerry G. Hofer, Nianbo Dong, Dale C. Farran, and Carol Bilbrey. 2013a. "Evaluation of the Tennessee Voluntary Prekindergarten Program: End of Pre-K Results from the Randomized Control Design." Research Report, Vanderbilt University, Peabody Research Institute, Nashville.

———. 2013b. "Evaluation of the Tennessee Voluntary Prekindergarten Program: Kindergarten and First Grade Follow-Up Results from the Randomized Control Design Research Report." Vanderbilt University, Peabody Research Institute, Nashville.

Ludwig, Jens, and Douglas L. Miller. 2007. "Does Head Start Improve Children's Life Chances? Evidence from a Regression Discontinuity Design." *Quarterly Journal of Economics* 122, no. 1: 159–208.

Masse, Leonard N., and Steven Barnett. 2002. "A Benefit-Cost Analysis of the Abecedarian Early Childhood Intervention." In *Cost-Effectiveness and Educational Policy,* edited by Henry M. Levin and Patrick J. McEwan. Routledge: Annual Yearbook of the American Education Finance Association.

McClelland, Robert, and Shannon Mok. 2012. "A Review of Recent Research on Labor Supply Elasticities." Working Paper no. 2012-12. Washington, D.C.: Congressional Budget Office.

Peisner-Feinberg, Ellen, Jennifer Schaaf, and Doré LaForett. 2013. *Children's Growth and Classroom Experiences in Georgia's Pre-K Program: Findings from the 2011–2012 Evaluation Study.* Chapel Hill: Frank Porter Graham Child Development Institute.

Puma, Mike, Stephen Bell, Ronna Cook, Camilla Heid, Pam Broene, Frank Jenkins, Andrew Mashburn, and Jason Downer. 2012. *Third Grade Follow-up to the Head Start Impact Study Final Report.* OPRE Report no. 2012-45b. Washington, D.C.: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.

Schweinhart, Lawrence J., Jeanne Montie, Zongping Xiang, William S. Barnett, and others. 2005. *Lifetime Effects: The HighScope Perry Preschool Study through Age 40.* Monograph no. 14 of the HighScope Educational Research Foundation. Ypsilanti, Mich.: HighScope Press.

Wolfers, Justin. 2006. "Did Unilateral Divorce Laws Raise Divorce Rates? A Reconciliation and New Results." *American Economic Review* 96, no. 5: 1802–20.

Wong, Vivian C., Thomas D. Cook, W. Steven Barnett, and Kwanghee Jung. 2008. "An Effectiveness-Based Evaluation of Five State Pre-Kindergarten Programs." *Journal of Policy Analysis and Management* 27, no. 1: 122–54.