DOES UNIVERSAL PRESCHOOL HIT THE TARGET? PROGRAM ACCESS AND
PRESCHOOL IMPACTS

Elizabeth U. Cascio

Does Universal Preschool Hit the Target? Program Access and Preschool Impacts
Elizabeth U. Cascio
NBER Working Paper No. 23215
March 2017, Revised July 2019
JEL No. H75,I24,I28,J13,J24

ABSTRACT

This paper studies the cost efficacy of universal over means-tested (targeted) programs, taking advantage of the rich diversity in state rules governing access to state-funded preschool in the U.S. Using age-eligibility rules for identification, I find that attending a state-funded universal preschool generates substantial immediate test score gains, particularly for poor children. Gains for poor children from attending targeted preschool are significantly smaller. These findings are robust to many specification checks, and cross-state differences in alternative care options, demographics, and other key features of the program environment cannot explain the difference in attendance impacts across program types. Impacts of universal public kindergarten and universal pre-K also look substantively similar within the same data, supporting an access interpretation. Benefit-cost ratios of universal programs are favorable despite their relatively high costs per poor child.

Elizabeth U. Cascio
Department of Economics
Dartmouth College
6106 Rockefeller Hall
Hanover, NH 03755
and NBER
elizabeth.u.cascio@dartmouth.edu

I.   Introduction

In the context of many public programs, key policy parameters involve not just how but *whom* – which populations should be eligible for benefits. This is evident in recent policy proposals at the federal level in the United States, which depart from the targeting characteristic of much U.S. policy to expand benefits to all individuals who meet broader eligibility criteria, even when many could be infra-marginal for the good or service in question. Central to the rhetoric behind such proposals, such as Medicare for All or universal child care, is that access to some services is a basic right, not a privilege for those who can afford it. Another rationale is that "programs for the poor are poor programs" – that means-tested programs end up underfunded due to lack of broad-based political support.

But can universal programs ever be justified on efficiency grounds? In theory, there are clear efficiency rationales for targeting: by targeting benefits on difficult-to-change characteristics correlated with low levels of human capital, policymakers can reduce moral hazard, keep costs down, and redistribute toward those most in need. On the other hand, the greater number – and political power – of stakeholders in universal programs might hold public goods providers more accountable, raising the productivity of public spending relative to a targeted program. Particularly in cases where there are direct interactions across program beneficiaries, universal access may also allow for human capital spillovers that increase the productivity of public spending for an implicit target population, such as disadvantaged children.

It is nevertheless difficult to gain empirical traction on this question; in the very least, it requires that the same kind of program be observed under widely different conditions of access, a situation that rarely arises in practice. Preschool education in the United States may provide a rare proving ground. Perhaps nowhere today is the variation in program access more striking: in

2015-16, not all states funded pre-kindergarten (pre-K) programs, and among the 43 states that did, there was great cross-state variation in eligibility rules. Some state programs are universal, serving all 4-year-olds that meet age-eligibility requirements; others are targeted, meaning that they are also means-tested or target enrollment based on other risk factors (Barnett et al., 2017).

In the first part of this paper, I take advantage of this rich cross-state variation in rules governing eligibility for state-funded pre-K to compare universal and targeted programs on the same basis, using the same data and research design. Despite a large literature on preschool education spanning disciplines, such an exercise has yet to be carried out. A mature body of research explores targeted preschool programs, most famously the federal Head Start program and the "model" preschool interventions of the distant past.[1] There is also emergent research on state-funded universal pre-K.[2] Both streams of literature tend to conclude that the benefits of preschool exceed the costs. Yet it is difficult to compare the findings for universal and targeted preschool programs directly due to differences across studies in methodology, outcomes, timing, and counterfactual enrollment patterns.[3]

I address this gap in the literature by working with survey data – the 2001 Birth Cohort of the Early Childhood Longitudinal Study (ECLS-B) – that span states where state-funded pre-K programs have different eligibility requirements and allow for credible estimation of the immediate gains from participation in universal and targeted programs alike. To estimate these

---

[1] Studies on Head Start include Currie and Thomas (1995), Garces, Thomas, and Currie (2002), Ludwig and Miller (2007), Deming (2009), Puma et al. (2010), Aizer and Cunha (2012), Carneiro and Ginja (2014), Bitler, Hoynes, and Domina (2014), Walters (2015), Kline and Walters (2016), Thompson (2018), Barr and Gibbs (2018), and Johnson and Jackson (forthcoming). Regarding "model" interventions, see Heckman et al. (2010), Schweinhart et al. (2005), and recent reviews by Elango et al. (2016) and Almond, Currie, and Duque (2017). There are also stand-alone studies on targeted state pre-K in North Carolina (Ladd, Muschkin, and Dodge, 2014) and Tennessee (Lipsey et al., 2013).

[2] See, for example, Gormley and Gayer (2005), Fitzpatrick (2008), Cascio and Schanzenbach (2013), and Weiland and Yoshikawa (2013).

[3] Wong et al. (2008) estimate the short-term cognitive effects of pre-K attendance in 2004-05 in five states. While the states differ some in terms of program access, the authors neither perform a formal analysis of the influence of access nor present estimates by family socio-economic status.

gains, I take advantage of the large differences in state-funded pre-K eligibility and attendance among 4-year-olds with birthdays near the threshold to be age-eligible for starting school. Specifically, I exploit the larger gap in pre-K attendance rates across adjacent school entry cohorts in states with more robust state-funded pre-K programs. I thus use a comparison group to account for unrestricted age and season of birth effects – a departure from the assumption of smoothness in a regression discontinuity (RD) model.[4] The rich background characteristics available in the ECLS-B, including pretests, also allow for useful tests of internal validity.

I find substantial positive effects of state-funded pre-K on the test scores of 4-year-olds in states with universal programs: universal pre-K eligibility (attendance) improves the average 4-year-old's standardized reading and math score of by a significant 12 percent (60 percent) of a standard deviation. I cannot rule out equal effects of universal pre-K attendance by family income, but low-income children experience substantially larger test score gains.[5] Pre-K attendance impacts for children in states with targeted programs are significantly smaller. Effect sizes vary, but this basic set of results is robust to changes in the estimation sample, including changing the range of ages or the set of states considered. It also arises for another outcome – parent reports of their 4-year-old's kindergarten readiness – albeit with less precision. Supporting a causal interpretation, my preferred specification also does not yield similar patterns of impacts for a measure of mental development at age 2 or a host of other child observables.

A key feature of the universal and targeted pre-K programs under study is that the per-pupil costs are actually quite similar. The larger test score gains for universal programs alone

---

[4] The pre-K evaluation literature has looked to age-eligibility thresholds as a source of identifying variation since Gormley and Gayer's (2005) pioneering application of the RD design to Tulsa's pre-K program. (See also, for example, Wong et al. (2008) and Weiland and Yoshikawa (2013).) However, these RD studies have relied on district or state administrative data on public school *students*, precluding consistent estimates of the impacts of pre-K eligibility and attendance impacts (Lipsey et al., 2015).

[5] I define low-income as eligibility for free- or reduced-price lunch, since this is the modal income requirement for targeted programs. The substantive conclusions are however robust to alternative definitions.

thus suggest they are more cost-effective. But this conclusion could be hasty since some other characteristic of states with universal programs – rather than program access itself – could generate the same pattern of attendance impacts. For example, universal programs were relatively more likely to require small class sizes, which have been rigorously shown to improve early test performance (Krueger, 1999). States with universal programs may also have populations more likely to benefit from formal preschool in general, or universal programs may be more likely to draw their enrollees from informal or parental care. I rigorously explore these possibilities in the second part of the paper and find little supporting evidence. I also show that, consistent with access as a causal mechanism, the impacts of universal pre-K look quantitatively similar to those of universal public *kindergarten* within the ECLS-B, which I estimate by exploiting age-eligibility rules for kindergarten entry using an RD approach.

The evidence is thus consistent with universal pre-K being relatively cost-effective. In the third and final part of the paper, I monetize the test score gains in a tentative cost-benefit analysis. Even under conservative assumptions on key parameters like the magnitude of the association between early life test scores and earnings, universal pre-K delivers a benefit-to-cost ratio well above one, much like universal kindergarten. While slightly less precise than the difference in test score effects of universal and targeted pre-K for poor children, the magnitude of the difference in benefit-to-cost ratios across programs is substantial.

The paper proceeds as follows. The next section describes the landscape of state-funded preschool in the U.S. and the state-funded pre-K programs of study in this paper. Section III outlines the research design I employ to estimate the causal impact of these programs on age 4 test scores, and Section IV introduces the data and presents an exploratory assessment of whether the design's identifying assumptions hold in the ECLS-B. Section V gives the main impact

estimates for universal and targeted pre-K, along with a series of specification checks on these estimates and their difference. Section VI explores potential alternative explanations for the larger impacts of pre-K attendance in states with universal programs beyond universal access *per se*, and Section VII offers the cost-benefit analysis. Section VIII concludes.

II.   Program Landscape

There has been striking growth in public funding of preschool programs since the early 1980s. Figure 1 shows trends from 1968 through 2015 in the number of states funding pre-K programs (left axis) and in enrollment rates of 3- and 4-year-olds in the federal Head Start program and in any public preschool (right axis).[6] In the early 1980s, only four states funded pre-K programs; by 2015-16, this figure reached 43 states and the District of Columbia. Public preschool enrollment rates have risen alongside this increased state funding commitment. This is particularly the case for 4-year-olds, for whom enrollment in Head Start – the other primary provider of public preschool – has stagnated since the early 1990s. State-funded pre-K programs indeed focus on 4-year-olds: during 2015-16, 32% of 4-year-olds were enrolled, compared to 5% of 3-year-olds (Barnett et al., 2017).[7]

The 2001 birth cohort of the ECLS-B – this study's focus – would have first aged into pre-K eligibility at age 4 in the fall of 2005, at the start of the 2005-06 school year. At this time (vertical line in Figure 1), state funding for pre-K was not that different than it's been more recently: 38 states and Washington, D.C. funded programs. As is also true in recent years, some programs had no eligibility requirements beyond age (universal programs), whereas others were

---

[6] Data on public preschool enrollment rates by age are calculated from the October Current Population Survey (CPS) School Enrollment supplements. Head Start enrollment rates divide Head Start enrollments reported by the Head Start Bureau by cohort size estimates based on Census Bureau estimates for July 1, 2005. State funding dates were constructed from program narratives published by NIEER (Barnett et al., 2017).

[7] The line between state-funded pre-K and Head Start is sometimes blurred, as some states allow school districts to subcontract with Head Start centers to provide pre-K. The pre-K enrollment measure used in this paper takes into account this possibility, as described to follow.

also means-tested or used other risk factors, like low parental education, to determine eligibility (targeted programs). My main analysis focuses on 16 state pre-K programs where in 2005-06, there as an enrollment differential favoring 4-year-olds of at least 8 percentage points and a state-established date by which the youngest enrollees were to have turned age 4 that did not fall in the middle of a month, according to statistics and program narratives published by the National Institute for Early Education Research (NIEER) (Barnett et al., 2006).[8]

Given these selection criteria, these 16 state pre-K programs were unsurprisingly among the larger ones operating in 2005-06. In fact, Figure 2 shows that of the five states with the largest pre-K programs in terms of age 4 enrollment shares at that time – Florida, Georgia, Oklahoma, Texas, and Vermont – only Vermont is excluded from the analysis (due to a having locally-determined entry cutoff birthdate). The population-weighted average state-funded pre-K enrollment rate for 4-year-olds (gap between age 4 and age 3 enrollment rates) across these 16 states was 34.3 percent (32 percentage points) in 2005-06, compared to only 9.5 percent (5.7 percentage points) in the remaining 22 states with programs.

Figure 2 also denotes which of these 16 states operated universal pre-K in 2005-06. The six universal states include Georgia and Oklahoma, which have the two longest-standing and most well-studied universal pre-K programs (Gormley and Gayer, 2005; Wong et al., 2008; Fitzpatrick, 2008, 2010; Cascio and Schanzenbach, 2013), whereas the targeted states include Tennessee, the only state to date with a pre-K program subjected to randomized evaluation (Lipsey et al., 2013). While greater political support for universal programs could translate into higher standards or higher per-pupil spending, this does not appear to have been the case at this time. Figure 2 shows no systematic relationship between universality and NIEER's 10-point

---

[8] I do not consider all states with pre-K programs due to constraints imposed by the ECLS-B and my empirical strategy. See Section III and Appendix A.

metric of suggested minimum state pre-K standards, and in fact targeted programs meet slightly more of these requirements on average (Appendix Table 1). Average per-pupil state spending on pre-K was also about the same in 2005-06 for the universal and targeted programs under study, at $3,500-$3,600 (nominal dollars, population weighted).[9]

If not via greater per-pupil resources, how might universal pre-K deliver larger benefits than targeted pre-K, especially for the poor children that both programs serve? Possible explanations include higher quality teachers[10] or higher academic expectations in universal pre-K classrooms,[11] peer effects, or perhaps a different mix of structural inputs for a given level of per-pupil spending. The universal programs under study actually on average imposed less teacher training than targeted ones, possibly in exchange for requiring smaller classes, lower staffing ratios, higher standards, and more comprehensive services (Appendix Table 1).

III. Empirical Strategy

I am interested in the achievement impacts from attending universal versus targeted state-funded pre-K programs. At base, the parameters of interest are thus the attendance impacts of each program type. Causal estimates are difficult to obtain: a simple difference in the scores of attendees and non-attendees will be a biased estimate of the attendance effect, since attendance is voluntary, and there might not be state funding to serve all children who are eligible. Children

---

[9] For most states, this figure represents only state contributions to pre-K. It therefore understates total spending, which could also be funded by local and federal revenue. I account for this possibility in the cost-benefit analysis of Section VII by taking K-12 per-pupil spending as an upper bound. K-12 spending is on average a bit higher in states with universal programs, at $11,875 per pupil versus $10,139 per pupil in states with targeted programs.

[10] Sabol et al. (2013) find that scores on the Classroom Assessment Scoring System (CLASS) do a better job than inputs (staff qualifications and class size) and learning environment (as measured by the Early Childhood Education Rating System – Revised, or ECERS-R) in predicting test score gains over the pre-K year. Exploiting random assignment of students to kindergarten teachers in a developing country, Araujo et al. (2016) find that kindergarten teachers with higher CLASS scores have higher value-added for reading and math scores. In a developing context, Araujo, Dormal, and Schady (forthcoming) also show that infants and toddlers quasi-randomly assigned to caregivers with higher CLASS scores have better fine motor, communication, and problem-solving skills.

[11] If prompted to focus on relatively advanced material, teachers may accelerate the learning gains of most students. For example, Engel, Claessens, and Finch (2013) show the more time teachers spend on more advanced mathematics content, the more children gain in math scores over kindergarten year, regardless of demographics.

who participate in state-funded pre-K may therefore have unobserved characteristics that directly influence their achievement.

I bring an empirical approach to this identification problem that is similar in spirit – but not implementation – to that taken first in the pre-K evaluation literature by Gormley and Gayer (2005). In particular, I wish to compare children with 4th birthdays near the threshold to be age-eligible for starting pre-K. Among children potentially served by a given pre-K program (all children for universal and the target group for targeted), those with 4th birthdays right on or before this birthdate threshold (e.g., Sept. 1) should have a much higher probability of pre-K participation in that school year than children with 4th birthdays right after. If children with birthdays near the age-eligibility threshold are on average similar along other dimensions that matter for test scores, comparing their scores later that school year should allow one to recover the causal effects of age-eligibility for pre-K, and scaling by the difference in their pre-K attendance rates, the causal effects of pre-K attendance.

In practice, the ECLS-B precludes such a sharp comparison; even if information on exact birthday were available (it is not), there would be too few observations on a daily basis to generate informative estimates. In past applications using administrative data from specific states or school districts, researchers have addressed this issue by considering a wider range of birthdates, but also recognizing that children with birthdays on opposite sides of the cutoff date no longer have the same potential on average. In fact, even if these children were to have similar unobservables, they differ along an observed dimension – age – that is strongly related to child development (e.g., Elder and Lubotsky, 2009). The RD solution is to assume that age effects on test scores are smooth or can be modeled with a polynomial function in age that is continuous through the birthdate cutoff.

The ECLS-B provides information on month of birth, not exact birthdate, and data on children across the U.S., not just in specific states or school districts. These data support an alternative, difference-in-differences (DD) approach – a comparison of the test scores of 4-year-olds in adjacent school entry cohorts in states with the state-funded pre-K programs identified in Section II (the treatment states), versus other states.[12] I work with 17 other, comparison states that had a state-established kindergarten entry cutoff birthdate in fall 2006 that was not in the middle of the month and either had no state-funded pre-K or state-funded pre-K enrollment rates that were too low or not different enough between 3- and 4-year-olds, according to NIEER (Appendix Table 2), for me to consider them treatment states.[13] Differences in test scores of 4-year-olds across adjacent school entry cohorts in the comparison states are intended to capture what would have happened for children in the treatment states in the absence of a state-funded pre-K program, due to aging or other factors.

Ignoring the distinction between universal and targeted programs for simplicity, the reduced-form DD model of interest that captures this idea is given by:

(1) $\qquad y_{is} = \theta elig_{is} \times treat_s + \sum_{m=-4}^{7} \gamma_m elig_{is}^m + \alpha_s + v_{is},$

where y_{is} is the age 4 (2005-06 academic year in the ECLS-B) test score of child i in state s, and $treat_s$ is a dummy equal to one if s is a treatment state. $elig_{is}$ is then a dummy equal to one if i is in the earlier (or older) entry cohort, set to enter kindergarten in fall 2006 rather than fall 2007 (and pre-K in fall 2005 rather than fall 2006 in treatment states). That is, $elig_{is} = 1[agek_i - agek_s^* \geq 0]$, where $agek_i$ is child i's age in months on September 1, 2006, and $agek_s^*$ is the minimum age

---

[12] In treatment states, cutoff birthdates for pre-K in fall 2005 are the same as those for kindergarten in fall 2006 (from Barnett, et al., 2007), so kids in adjacent pre-K entry cohorts are also in adjacent kindergarten entry cohorts.
[13] The comparison states thus include states without pre-K programs and states whose pre-K programs did not serve 4-year-olds nearly exclusively (see Figure 2). For the second group of states, it would be impossible to detect a first stage using this empirical approach and the ECLS-B. All comparison states are listed in Appendix Table 2, and the process of their selection is described in Appendix A.

in months for kindergarten entry in state s on September 1, 2006.[14] Thus, if all states had September 1 cutoff birthdates, $elig_{is}$ would equal one for all ECLS-B respondents born January through August 2001 and zero for those born September through December 2001.[15] The $elig_{is}{}^m$ then represent a series of dummies for age in months relative to the minimum age in months for kindergarten entry, or $elig_{is}{}^m = 1[agek_i - agek_s{}^* = m]$, where $-4 \leq m \leq 7$.[16]

The coefficient of interest in model 1 is the intent-to-treat (ITT) effect θ, which captures how much more entry cohort, or pre-K eligibility, relates to the age 4 (2005-06 academic year) test scores of children in treatment states. But recall that the parameter of interest is the impact of state-funded pre-K *attendance*, not eligibility. With data from the ECLS-B, I am able to produce estimates of the effects of the treatment on the treated (TOT) by instrumenting for state-funded pre-K attendance, $prek_{is}$, in the model

(2) $y_{is} = \beta prek_{is} + \sum_{m=-4}^{7} \lambda_m elig_{is}^m + \delta_s + \varepsilon_{is}$

with $elig_{is}$ x $treat_s$, using two-stage least squares (TSLS).

For TSLS estimation of (2) to produce unbiased estimates of β, it must be the case that differences in unobserved determinants of outcomes across entry cohorts do not systematically differ between the treatment and comparison states, or $elig_{is}$ x $treat_s$ is uncorrelated with ε_{is} conditional on the $elig^m{}_{is}$ and state fixed effects. This identifying assumption is weaker than in the RD case, since the direct effects of the running variable, $agek_i - agek_s{}^*$, are not modeled as

---

[14] Without exact birthday, I must include children who are eligible for kindergarten in fall 2006 (e.g., those turning 5 on September 1, 2006 in a state with a September 1 cutoff date) in the fall 2007 kindergarten cohort. If births are uniformly distributed across days, this should lead to a small attenuation bias in estimates of the eligibility impacts. Thus, I minimize attenuation bias by excluding states with cutoff dates closer to the middle of the month.

[15] Ten of the 16 treatment states and 12 of the 17 comparison states actually do require entering pre-kindergartners to be age 4 on August 31 or September 1 – roughly the start of the typical school year. Many of the remaining states (3 of the treatment states and 1 comparison state) require pre-kindergartners to be age 4 one month later (September 30 or October 1). See Appendix Tables 1 and 2.

[16] Because September 1 is the modal cutoff date (Appendix Tables 1 and 2), I restrict the sample to children born in the 8 months before their state cutoff date ($0 \leq agek_i - agek_s{}^* \leq 7$) or in the 4 months after ($-4 \leq agek_i - agek_s{}^* \leq -1$).

smooth or continuous through the age-eligibility threshold. Identification also requires a significant first stage, or that $elig_{is}$ x $treat_s$ significantly predicts pre-K attendance. The first-stage coefficients on the instrument will by (program) design differ in models that do not distinguish respondents by family income, making TSLS estimation of model 2 critical to comparing the benefits of universal and targeted programs. I begin by pooling respondents across the family income distribution. However, I quickly move to presenting estimates by family income, since targeted programs should only affect lower-income children.[17]

Being able to estimate TOT impacts is one advantage of using the ECLS-B, as the administrative data used in previous RD applications have been restricted to students who enroll in public pre-K (Lipsey et al., 2015). But there are other advantages to these survey data. For example, the ECLS-B provides a rich set of baseline characteristics, including birth weight and earlier (age 2) outcomes, with which to evaluate the internal validity of the empirical approach, and data on alternative care and education options, allowing me to better understand the counterfactual to the state-funded program. The age 4 (2005-06) ECLS-B test is also designed to be age-appropriate and is administered before children attending pre-K would have progressed to kindergarten, limiting contamination from kindergarten exposure possibly present in RD studies.

IV.   Data and Exploratory Analysis

As discussed, the validity of my empirical approach rests on two assumptions. The first is that there is a first-stage relationship between eligibility and state-funded pre-K attendance, or that the interaction of school entry cohort and residence in a treated state, $elig_{is}$ x $treat_s$, predicts state-funded pre-K attendance in the 2001 birth cohort as of the 2005-06 academic year. The

---

[17] In a departure from an earlier draft of this paper, I no longer begin with stratification by family income for two reasons. First, the cost-benefit analysis for universal programs must incorporate higher-income children. Second, I cannot perfectly replicate the other eligibility rules for targeted programs in the ECLS-B; the low-income indicator that I rely upon (indicating eligibility for free- or reduced-price school lunch) is the best available measure.

second is that this interaction does not predict unobserved correlates of test scores. In this section, I provide some preliminary evidence that these assumptions are met and provide details on the construction of key variables and the estimation sample from the ECLS-B.

A.   State-funded Pre-K Attendance

Administrative information on whether the ECLS-B respondents attended state-funded pre-K is unavailable. However, the survey provides detailed information on the care and education of respondents at age four (wave 3 of the survey, corresponding to 2005-06) from interviews with both parents and providers. While interviews of providers could in principle offer more reliable information, they were only administered to the one program, center, or person accounting for the *most* of a given child's non-parental care, and thus fail to pick up all cases of enrollment in state-funded pre-K. My base measure of pre-K attendance is therefore the parent report of a child attending free pre-K or free preschool. A limitation of this measure alone, however, is that parents whose children attended a state-funded pre-K program delivered via Head Start may have reported Head Start participation instead. In these cases – and in fact in all cases where the provider report signals public pre-K attendance when the parent report does not – I adjust my pre-K attendance variable upward accordingly. I code state-funded pre-K attendance in the provider interviews as public-school pre-K or another program (e.g., preschool or child care) sponsored by state or local government or school district. (See Appendix A.)

The first graph in Figure 3 Panel A shows pre-K attendance rates in 2005-06 by program type in one's state of residence (universal, targeted, comparison) and age relative to the minimum age for entering kindergarten the subsequent school year (in two-month intervals to reduce noise). Age is increasing along the horizontal axis, with the first two points representing ages of children who would not have been eligible for kindergarten in fall 2006 (or pre-K in fall

2005). The second graph in Figure 3 Panel A then shows the difference in means between each

of the two treatment groups and the comparison group, relative to what that difference was for

children who just missed eligibility ($-2 \leq agek_i - agek_s^* \leq -1$), adjusting for state and month of

assessment fixed effects. Aside from age being grouped into intervals, this is a generalization of

model 1, with pre-K attendance as the outcome. For now, the estimation sample includes *all*

children, regardless of family income, who were resident in one of the 16 treatment or 17

comparison states in wave 3 of the ECLS-B and were 5 years old between 8 months before and 4

months after their state's kindergarten entry cutoff.[18]

Pooling across family income, the first-stage impacts of age-eligibility alone should be

lower in states with targeted pre-K programs, where eligibility is also based on family income or

other risk factors. This expectation is realized in the data. As shown in Panel A, all three groups

of states exhibit similar, relatively low 2005-06 pre-K attendance rates among children who were

not eligible to attend kindergarten in fall 2006 (or pre-K in 2005 in treatment states). However,

pre-K attendance rates between treatment and comparison states diverge for children age-eligible

to start pre-K in fall 2005, and the extent of divergence is greater for children in universal

programs, as anticipated. As shown in the second graph, the regression-adjusted DD estimates

for these age-eligible children are statistically significant.

Columns 1 and 3 in Table 1 Panel A give mean pre-K attendance rates among age-

ineligible children in treatment states (i.e., eligible for pre-K in fall 2006, not fall 2005).

Columns 2 and 4 show the first-stage DD estimates that correspond to the second subpanel in

---

[18] I additionally restrict attention to children with non-missing preschool-age cognitive assessments administered in
September or later and non-missing demographic and background characteristics. These additional restrictions lead
to a loss of very few observations. There are 5,100 observations in the sample overall; 1,750 of these children reside
in states with targeted programs; 1,150 reside in states with universal programs; and 2,250 reside in the comparison
states. Reported sample sizes are rounded to the nearest 50, per IES rules to protect confidentiality of ECLS-B
respondents.

Figure 3 Panel A, i.e., subgroup-specific coefficients (standard errors) on $elig_{is}$ x $treat_s$ from

model 1, which includes fixed effects for state of residence and single months of age relative to

the threshold, with $prek_{is}$ as the dependent variable.[19] The estimated gap in 2005-06 pre-K

enrollment rates between the 2006 and 2007 kindergarten cohorts was 21.1 percentage points

higher in states with universal pre-K programs than in the comparison states (column 2). For

targeted states, on the other hand, this gap amounted to 11.4 percentage points (column 4). The

difference in these estimates is statistically significant (column 5).[20]

B.   Demographic and Background Characteristics

       Table 1 Panel B gives ineligible means and analogous DD estimates for demographic and

background characteristics in the full sample. In addition to basic demographics – age at

assessment (in months) and indicators for sex (female) and race (non-Hispanic black and

Hispanic) – I construct indicators for low birth weight (birth weight < 2,500 grams), for low

maternal education (at or below a high school degree), for a language other than English being

spoken in the home, for the presence of both biological parents in the household, and for low

family income. I define low-income as eligibility for free or reduced-price lunch (family income

≤ 185% of the federal poverty line (FPL)), since it is the modal eligibility criterion for the

targeted states under consideration and thus the best available way to stratify the analysis.[21]

       If the TSLS estimates are identified, $elig_{is}$ x $treat_s$ should have little predictive power

---

[19] As in Figure 3, I cluster standard errors on state of residence-by-month of birth (the level of the treatment) and weight the analysis using sampling weights appropriate for analyses using data from the first and third waves of the ECLS-B. The model also includes dummies for month-by-year of the wave 3 assessment. Reported sample sizes are rounded to the nearest 50, per IES rules to protect confidentiality of ECLS-B respondents.

[20] Reassuringly, the universal-targeted gap in first-stage coefficient estimates (21.1/11.4 =1.85) is proportionally similar to the universal-targeted gap in the difference in age 4 and age 3 state funded pre-K participation rates reported by NIEER for 2005-06 (41.2/23.1=1.78).

[21] The eligibility criterion is relevant for five targeted states – Texas, Maryland, Louisiana, Colorado, and Tennessee – which together account for 56 (69) percent of 4-year old population (state-funded pre-K enrollment) in the states with targeted programs listed in Appendix Table 1. The remaining states have different income requirements (Michigan, Kansas) or no income requirements, but risk factor requirements that correlate strongly with income (Illinois, South Carolina, Virginia). I test the sensitivity of my conclusions to the definition of poverty below.

with regard to these observed correlates of test scores, just as it should have little predictive power for unobservables. For the most part, the coefficient estimates in the even columns are not statistically significant. There are exceptions, however. Most notably, children who are age-eligible for universal preschool are significantly less likely to be poor; this leads me to reject the test of joint significance on the DD coefficients (including this variable, the *p*-value on joint test is 0.01; excluding it, it is 0.32; column 2).[22] This phenomenon is also evident, but to a lesser extent, among children in targeted states (column 4); for both subsamples, the DD coefficient for age in months is also negative and (marginally) statistically significant, though small.

These findings, and the fact that some of these coefficients are large even if not statistically significant, suggest the importance of including these background variables as controls in the analysis. They also suggest the importance of seeking additional ways to validate the research design. I do so below in considering both prior (age 2) cognitive test scores as an outcome, and in testing for "impacts" of eligibility among ineligible students, the idea being that such analyses will only turn up significant coefficients if there is confounding by unobservables. It is important to note, however, that poverty and age are balanced for the *difference* in estimates between universal and targeted states, as demonstrated by the triple difference (DDD) estimates in column 5, and it is the difference in program effects that is the focal point of the paper.

V.   Effects of Pre-K Eligibility on Preschool-Age Test Scores

A.   Baseline Estimates

I focus the analysis of outcomes on cognitive test scores from the third (preschool-age) wave of the ECLS-B. The preschool-age cognitive assessment included math and reading components and was designed to test both for developmental (age-based) milestones and for

---

[22] This is more likely an artifact of sampling variation than of explicit sorting (Dickert-Conlin and Elder, 2010).

knowledge and skills considered important for school readiness and early school success (see Appendix A). I standardize test scores to have a mean of zero and a standard deviation of one in the comparison states, and calculate average scores across reading and math as an unweighted mean of these standardized scores. Test administration was concentrated during the fall of the 2005-06 school year, however, possibly raising concerns that not enough time would have elapsed for children's scores to reflect pre-K. In my robustness analysis, I therefore consider an alternative outcome at preschool-age – an indicator for whether a parent reports concern over a child's readiness for kindergarten. This measure captures variation across program types in parental perceptions of likely impacts of pre-K attendance, not just on academic preparation for kindergarten but also behavioral and social preparation as well.[23]

Mean preschool-age test scores for the ineligible subsample in treatment states, shown in column 2 of Table 2, are negative not because universal or targeted states are negatively selected, but rather because there is a strong age gradient in test scores in all states. This is evident in the first graph in Figure 3 Panel B: for comparison states, where the age gradient is least likely to reflect pre-K by design, average reading and math test scores rise about 80% of a standard deviation between the youngest and oldest children in the estimation sample. There is thus significant variation in early test performance based solely on age, holding constant family background. This age gradient in test scores provides a useful benchmark to which to compare the estimated impacts of pre-K attendance.

The remaining curves in Figure 3 Panel B show how the existence of a robust state-funded pre-K program affects this age gradient in preschool-age test scores. There is a clear and relatively sustained divergence between the average test performance of age-eligible children in

---

[23] The ECLS-B provides one socio-emotional assessment at age 4 – the "two bags task." I analyzed this assessment in an earlier version of this paper, but the estimates were uninformative.

universal pre-K states relative to the comparison group (subpanel 1). The relative gains in test performance are statistically significant for the two youngest groups of children in the eligible cohort (subpanel 2). While there is some suggestion of such gains for targeted programs, the effect dies out more quickly, and even turns negative for the oldest children in the sample. Combined with the evidence for pre-K attendance in Panel A, these figures are consistent with children gaining more in the short-term from pre-K attendance in states with universal programs. Supporting this inference, there is no pre-existing upward trend in the test score gap between treatment and comparison states.

Table 2 presents the corresponding first-stage and reduced-form (model 1) estimates of the impacts of pre-K eligibility and IV and OLS (model 2) estimates of the impacts of pre-K attendance on preschool-age test scores. I show estimates without and then with controls for the demographic and background characteristics in Table 1 Panel B, separately for universal (Panel A) and targeted (Panel B) programs. Though the additional controls have little impact on first-stage estimates, they more moderately impact the reduced-form and IV estimates. With controls I reject the null that attending a targeted pre-K program has the same effect as attending a universal pre-K program, as shown in Panel C (p=0.046). The IV estimated impact of attending pre-K in a universal state also remains substantial, at a marginally statistically significant 0.57 standard deviations – about 70% of the impact that the average child would expect from a full year of aging. This estimate is significantly different from its OLS counterpart (-0.07 standard deviations), suggesting negative selection into pre-K attendance.

B.   *Heterogeneity by Family Income*

I prefer estimates from a sample including children regardless of family income as a starting point, since I cannot perfectly replicate the other eligibility rules for targeted programs in

the ECLS-B. Table 3 presents estimates applying arguably the best available consistent

definition, which cuts the sample by eligibility for free- or reduced-price lunch, based on the

preferred specification with the additional controls; group-specific means and balance tests are

provided in Appendix Table 3. This cut of the data produces income gaps in achievement like

those that have been seen in other data (Table 3 column 2) (Reardon 2011). First-stage impacts

of age-eligibility for pre-K (column 1) are also essentially the same for poor children regardless

of whether a program is universal or targeted (0.226 versus 0.223) but are only evident for other

children in states with universal programs (0.189 versus 0.037), as expected. Appendix Figure 1

provides transparent graphical evidence of these first-stage impacts.

For poor children, estimated test score gains from pre-K are strikingly different for states

with universal versus targeted programs. Indeed, although it remains possible that universal pre-

K (Panel A) leads to equivalent gains across the two family-income groups (p=0.118 for the

reduced-form and p=0.191 for IV),[24] the estimates for universal programs shown in Table 2 and

Figure 3 were clearly driven by the low-income subsample. Age-eligibility for pre-K raises the

preschool-age test scores of poor children in universal states by a significant 0.26 standard

deviations (Panel A, column 3). By contrast, it reduces the test scores of poor children in targeted

states by an insignificant 0.018 standard deviations (p=0.02 on the difference; Panel C). Figure 4

suggests that this basic insight also holds when I divide the sample by quintiles of a socio-

economic status (SES) index provided by the ECLS-B, derived from factor analysis on parental

education, parental occupation, and family income: test score gains from pre-K in universal

states are concentrated in the two bottom quartiles of this index (Panel B) despite positive

---

[24] Finding larger effects of pre-K attendance for poor children would be consistent with a framework where higher-income children have relatively high-quality care and education options in the absence of universal pre-K and with much existing evidence on universal preschools both in the U.S. and worldwide (Cascio, 2015; Elango et al., 2016).

impacts on pre-K attendance across the income distribution (Panel A).

Returning to Table 3, the IV estimate implies that pre-K attendance raises the preschool-aged test scores of poor children in universal states by 1.16 standard deviations. While such an estimate might see quite large, it is in line with recent estimates of the effects of the earliest years of universal elementary education (e.g., Anderson et al., 2011; Fitzpatrick, Grissmer, and Hastedt, 2011).[25] For these children, pre-K attendance also improves performance substantially more than aging alone would: as shown in the plot of means in Figure 5, the age gradient in test scores is relatively shallow for low-income students in the comparison group, presumably because of their relatively weak cognitive stimulation in the absence of state-funded pre-K. Evidence presented to follow confirms that low-income children are substantially less likely to experience formal education or care in the absence of state-funded pre-K.

Figure 5 presents graphs the test score impacts of pre-K by family income in an analogous way to Figure 3 Panel B. The concentration of positive impacts of universal pre-K among the youngest eligible students is now even more evident (Panel A). Separate analyses of the subcomponents of the test (Appendix Figures 2 and 3) reveals that performance on the math subcomponent is responsible for this pattern; effects on reading scores are more sustained. Moreover, estimating the preferred model separately for the standardized reading and math scores, as done in Panels B and C of Table 4 for the full sample (Appendix Table 4 for the low-income subsample), reveals universal pre-K effects that are much more precisely estimated for reading scores. Pre-existing trends in reading scores are also relatively similar across the three groups of states (Appendix Figure 2), lending greater credibility to those findings.

---

[25] Likewise, though imprecise, the smaller effect size for a poor child attending pre-K in a targeted state is consistent with the relatively small short-term test score gains from targeted preschool from recent randomized evaluations (Puma et al., 2010; Kline and Walters, 2016; Lipsey et al., 2013).

The estimates presented thus far suggest that children attending pre-K in states with universal programs experience larger early test score gains than children attending pre-K in states with targeted programs, and the difference is more pronounced for the low-income children who would likely meet eligibility criteria for either type of program. But these findings could also be an artifact of the design, estimation sample, or choice of outcome. In this section, I assess the robustness of the basic set of results to these decisions.

In Table 4 Panel D, I first consider impacts of pre-K attendance on a test of mental development at age 2 (or in wave 2 of the ECLS-B), before children would have been eligible for pre-K.[26] Performance on this test is neither significantly affected by pre-K attendance nor significantly different across states with universal and targeted programs ($p=0.613$), suggesting limited contamination by unobservables. The estimates for age 4 test impacts are also similar when age 2 scores are included as controls (Appendix Table 5).[27]

In Table 4 Panel E, I then consider an alternative outcome at age 4 – an indicator for whether a parent reports that their child is not ready for kindergarten. Mirroring the estimates for test scores, the parents of children attending pre-K in states with universal programs are less likely to express concern over their kindergarten readiness. The estimates are more pronounced among lower-income parents (Appendix Table 4), but as was also the case with test scores, I cannot rule out that they are identical across income groups ($p=0.35$). Unlike in the test score

---

[26] The wave 2 cognitive assessment is the Bayley Short Form-Research Edition (based on the Bayley Scales of Infant Development, 2nd Edition). Results with Bayley motor scores as an outcome show a similar pattern as those for mental scores and so are omitted for brevity. The model in Panel D also includes age of assessment and dummies for month-by-year of assessment in the relevant wave, in addition to the controls of the baseline model, and is weighted by sampling weights appropriate for inclusion of that wave.

[27] Controlling for age 2 scores in the full sample has no appreciable effect on the estimates for universal programs but raises estimates for targeted programs to an extent that the gap in IV estimates across program types is no longer statistically significant ($p=0.178$). The difference in IV estimates for poor children does, however, remain marginally significant with these controls included ($p=0.055$).

case, however, I also cannot rule out that the estimates are the same across program types for poor children ($p=0.22$), though the evidence is slightly more suggestive for a reduction in concern over *academic* readiness for kindergarten in particular ($p=0.17$; Appendix Table 6). Similar findings do not arise for parent reports of concern over school readiness on other margins (Appendix Table 6), suggesting that the relative benefits of universal pre-K may be primarily academic. But the relative lack of statistical power makes all of these complementary estimates suggestive at best.

I consider several changes to the estimation sample in Table 5, returning to the original test score outcome. I included a relatively wide age span for eligible children in the baseline estimation sample to improve statistical power, but possibly at the expense of greater bias. Limiting estimation to children with birthdays within 4 months of the age-eligibility threshold (Panel B), who are more similar to one another along some but not all dimensions (Appendix Table 7), actually generates more positive estimates for the impacts of pre-K eligibility and attendance, enough so that I am less confident that the IV estimates differ between universal and targeted states in the full sample ($p=0.197$). However, I remain reasonably confident in this conclusion for low-income children, even in this restricted sample ($p=0.060$; Appendix Table 8).

The two remaining changes to the estimation sample represented in Table 5 make it more, then less, expansive. Adding treatment states may further help limit the influence of idiosyncratic state samples, since the ECLS-B is not designed to be state-representative. With this in mind, I expand the sample to include children in the two targeted states (Arkansas and North Carolina) and one universal state (Maine) with middle of the month birthdate cutoffs, as long as their birthdate is not in the cutoff month (to minimize misclassification), and results are not much changed (Panel C). On the other hand, limiting attention to the 14 treatment states (5

universal, 9 targeted) and 15 comparison states with cutoff birthdates on or before October 1 –
with the goal of keeping the state composition of the sample of ineligible students more constant
across the age span – lowers the IV estimate to 0.52 standard deviations. However, state-funded
pre-K attendance in universal states still yields larger test score gains (Panel D).

VI.   Interpretation

It may be tempting to conclude that universal access is itself the driver of the generally
robust finding of larger effects of pre-K attendance in states with universal programs. However,
whether a pre-K program is universal or not is not randomly assigned; state contexts, pre-K
programs and populations differ along other dimensions. In this section, I attempt to rule out
leading alternative explanations – differences in counterfactual care, other pre-K characteristics
(specifically maximum class size), and demographics in states with universal programs. I also
present additional evidence consistent with an access interpretation: the estimates for universal
pre-K look much like what one finds for universal *kindergarten*, at age 5, within the ECLS-B.

A.   Differences in the Counterfactual?

If universal pre-K attendees were drawn less from other center-based care and more from
informal or parental care, we would expect to see relatively large test score effects of universal
pre-K, all else constant.[28] I estimate the reduced-form impact of age-eligibility for pre-K and the
IV effect of pre-K attendance for a mutually-exclusive set of alternative care options – Head
Start, other center-based care, informal non-parental care, and parental care.[29] For additional
statistical power, I also combine Head Start and other center-based care as "formal care" and

---

[28] In re-analyses of data from the Head Start Impact Study, Feller et al. (2016) and Kline and Walters (2016) find
that Head Start has much smaller impacts on children who would have otherwise been in center-based care.
[29] The reduced-form DD coefficients therefore add up to zero across all categories, including the first-stage DD
coefficient for pre-K. Reassuringly, reported Head Start enrollment rates for low-income children in universal and
targeted states (at 12%) are quite similar to what they are in the administrative data.

informal non-parental and parental care as "informal care."

The IV estimates for the full sample (Appendix Table 9) imply that approximately 90% of targeted pre-K attendees were drawn from informal care, versus 46% of universal pre-K attendees. The difference is not surprising: universal programs serve higher-income children, for whom formal care or education in the absence of state-funded pre-K is relatively common. More telling are findings for the low-income subsample, where the difference in estimated attendance effects across program types is most pronounced. Table 6 shows substitution patterns that are more similar in this sample, with point estimates implying that about 35% of poor pre-K attendees would have otherwise attended Head Start and about 15% would have otherwise been in informal, non-parental care. But parental care was a more likely alternative in the targeted case, so that 70% of poor targeted pre-K attendees would have otherwise been in informal care, compared to 55% of poor universal pre-K attendees.[30] Though the difference is not statistically significant, this suggests that the estimated gap in pre-K attendance effects across program types is actually *lower* than it would be with a constant counterfactual. It therefore does not provide strong evidence against the conclusion that universal programs outperform targeted ones.

B.   *Differences in Other Characteristics of Programs or Populations?*

As noted in Section II, though universal and targeted programs look similar in the aggregate in terms of available resources and the overall number of minimum standards to which they are held, the program types on average differ in how funds are allocated or which standards are emphasized, and it might be these differences in resource allocation that are driving the differential in estimated effects across states with different pre-K access. Of particular interest is that the universal programs under study prioritize smaller class sizes, which in early education

---

[30] These figures for program substitution are in the same ballpark as found in the recent Head Start Impact Study (Feller et al., 2016; Kline and Walters, 2016).

have been convincingly shown to produce higher immediate test scores (Krueger, 1999). With

the important caveat that the decision to require small class sizes could be endogenous to the

access decision itself, I can investigate whether the difference in reduced-form and IV estimates

across universal and targeted states is robust to allowing for heterogeneity in pre-K impacts by

this program dimension. Using a similar approach, I also explore robustness to regression-

adjusting for heterogeneity in pre-K impacts by the demographic and background characteristics

in Table 1 Panel B.[31]

Table 7 Panel A presents the difference in reduced-form (columns 1 and 4) and IV

(columns 2 and 5) estimates across universal and targeted programs, along with estimates of the

coefficients capturing how class size standards influence the effect of pre-K eligibility (columns

3 and 6), both for the full sample and for the poor subsample.[32] The universal-targeted difference

in reduced-form estimates shrinks somewhat, as pre-K eligibility has a greater impact on test

scores when class sizes are required to be small, as anticipated. But the difference in IV estimates

remains at least as large in both samples as it was at baseline. The estimates in Table 7 Panel B

similarly show demographic heterogeneity in the impacts of pre-K exposure in the expected

---

[31] Specifically, I begin with the triple-difference reduced-form model:
$$y_{isp} = \theta_{UT} elig_{is} \times treat_s \times uni_s + \theta_T elig_{is} \times treat_s + \theta_D elig_{is} \times treat_s \times C_s + \delta elig_{is} \times C_s +$$
$$\sum_{p=0}^{1} \sum_{m=-4}^{7} \gamma_{mp} elig_{is}^{m} \times P_p + \alpha_{sp} + \sum_{p=0}^{1} (x_{is} \times P_p)' \pi_p + \omega_{isp},$$
where uni_s and $treat_s$ represent, respectively, a dummy for whether s is a universal (treated) state and a dummy for whether s is a treatment state at all (universal *or* targeted), and C_s the number of class size standards (out of two possible) that the state program requires. All of the controls are interacted with an indicator for the estimation sample, $P_j=1[j=p]$, with $p=1$ for the universal pre-K estimation sample and $p=0$ for targeted pre-K estimation sample. Excluding the interactions with the class size variable ($elig_{is}$ x $treat_s$ x C_s and $elig_{is}$ x C_s), θ_{UT} is then the difference in the reduced-form DD estimates for universal and targeted pre-K programs presented in Table 2, and θ_T is the baseline reduced-form DD estimate for targeted programs. Of interest is then estimation with the class size interactions included, which adjusts estimates of θ_{UT} for the correlation between a state program offering universal access and the program requiring small classes. The coefficients on *prek_is* and *prek_is* x *uni_s* in the structural model of interest are estimated using TSLS with $elig_{is}$ x $treat_s$ x uni_s and $elig_{is}$ x $treat_s$ as excluded instruments. The model with demographic heterogeneity replaces C_s with X_i, where X_i is one of the characteristics in Table 1 Panel B. I also estimate version of this model including all of these characteristics simultaneously.
[32] With reference to the reduced-form model presented in the prior footnote, columns 1 and 4 present estimates of θ_{UT}, and columns 3 and 6 present estimates of θ_D.

direction: effects of pre-K eligibility tend to be larger (and are often statistically significant) for more disadvantaged populations, like children of mothers with no more than a high school degree. However, the populations of universal and targeted states are similar enough that this has no large impact on the conclusion that universal programs outperform targeted ones.

C.   Comparison to the Impacts of Kindergarten Using an Alternative Approach

To provide further evidence consistent with an access interpretation, I examine whether the test score impacts from attending universal public *kindergarten* look substantively similar to those from attending universal pre-K in the ECLS-B.[33] The fourth (kindergarten-age) wave of the ECLS-B also included math and reading cognitive assessments, again largely administered in the fall of the academic year, which I standardized and aggregated in a similar fashion to the preschool-assessments. It also included information on grade of enrollment and whether the school was public or private (if applicable).

To carry out this analysis, however, I must rely on a different (and arguably less robust) empirical approach than that earlier in the paper. In particular, since universal kindergarten is available in public schools across all states, I can only exploit age-eligibility rules using an RD model in the reduced-form, given by:

(3) $y_{is} = \tilde{\theta} elig_{is} + f\left(normage_{is}; \tilde{\lambda}\right) + \tilde{\alpha}_s + \varpi_{is}.$

$f\left(normage_{is}; \tilde{\lambda}\right)$ represents some smooth function, with parameter vector $\tilde{\lambda}$, of the difference between child i's age in months on September 1, 2006 and the minimum required age in state *s* on that date for kindergarten entry, i.e., *normage$_{is}$=agek$_i$ - agek$_s$**. For simplicity and because it is a good fit to the data (Appendix Figure 4), I specify the smooth function in normalized age as

---

[33] As earlier noted, the magnitude of the impacts of universal pre-K attendance in the preferred model appears comparable to that of universal early education more generally, but the relevant studies (Anderson, et al., 2011; Fitzpatrick, Grissmer, and Hastedt, 2011) used a different data source, the ECLS-K.

linear within +/- four months of the entry age threshold, with a different slope among those eligible for public kindergarten. I produce estimates of the impacts of K attendance by substituting a smooth function of the same form as in model 3 for the unrestricted eligibility effects in model 2, then estimate that model using TSLS with $elig_{is}$ as an instrument for $prek_{is}$.

Estimates for the full sample of states considered in the main analysis are given in Table 8. Underlying models also include as controls the demographic and background variables in Table 1 Panel B, as well as age and dummies for month-by-year of assessment during the wave 4 interview. There is a strong first-stage relationship between age-eligibility for kindergarten and public kindergarten attendance in 2006-07: on average, a child just barely eligible was 66 percentage points more likely to be enrolled in the full sample (column 1).[34] The marginal child in the full sample also scored a significant 0.38 standard deviations higher on the kindergarten-age reading and math tests (column 2). Point estimates are much more similar across income in this case, and the implied effect of public kindergarten attendance in the sample overall – 57% of a standard deviation (column 3) – is very similar to what I found for universal pre-K (Table 2).[35]

VII. Cost-Benefit Analysis

As laid out in Section II, per-pupil costs of the universal and targeted pre-K programs under study are similar. However, the attendance impact for the average child eligible to attend a universal program – approximately 0.6 standard deviations (Table 2) – greatly exceeds the attendance impact for the average (poor) child eligible for a targeted program – approximately -0.08 standard deviations (Table 3). While the findings thus suggest a more favorable benefit-to-

---

[34] Notably, this estimate is considerably lower than the *average* kindergarten enrollment rate of age-eligible children. This is to be expected given the commonly-found extent of non-compliance with entry-age regulations, but lack of information on exact birthday in the ECLS-B may attenuate this estimate further. I also find a low first-stage coefficient on the instrument for pre-K attendance relative to the levels of pre-K attendance reported by NIEER.
[35] However, substitution from alternative care is also more common than in the universal pre-K case: 73% switch from other formal center-based care (Appendix Table 10).

cost ratio for universal programs, it is useful to formalize this calculation.

Table 9 Panel A presents estimates of the benefit-to-cost ratio for each program type.[36] Only per-pupil state outlays are available, so I make two conservative assumptions regarding total per-pupil spending that are well above the state contribution. The marginal social benefit of universal pre-K attendance, measured as the present discounted value of the expected earnings gains from the increase in test scores, is 39% higher than the per-pupil cost of K-12 schooling (Panel A, row 1). It is 85% higher than the per-pupil cost of K-12 schooling net of the social savings from substitution from other public and private centers. These ratios are not statistically greater different than one but are close to significantly different from those for targeted pre-K programs (column 3), as seen in column 4. Using Head Start to approximate program outlays (Panel A row 2) considerably increases these ratios for universal pre-K. The 3.52 benefit-to-cost ratio in the most favorable (yet still conservative) scenario for universal pre-K is both greater than one and greater than the corresponding ratio for targeted programs. It is also larger than the (highly significant) benchmark estimate for kindergarten programs, of 2.96 (column 1).

Through it is customary to present benefit-to-cost ratios for early educational programs, the marginal value of public funds (MVPF) – or the ratio of a program beneficiary's willingness to pay for the program out of his/her own income to government costs net of fiscal externalities – provides a means of using causal estimates from program evaluation for social welfare analysis (Hendren, 2016). I thus also present estimates of the MVPF of each program, making a small adaptation to the MVPF formula used by Kline and Walters (2016) in their re-evaluation of the Head Start Impact Study to accommodate the contemporaneous private benefits (to families) of universal programs. Specifically, the MVPF in the present context pins the marginal program

---

[36] See table notes and Appendix B for a description of assumptions over key parameter values.

beneficiary's willingness to pay to both their net-of-tax earnings gains from program

participation and the reduction in their family's immediate out-of-pocket childcare expenses.

In particular, for program j, the MVPF is given by

$$MVPF^j = \frac{(1-\tau)p\hat{\beta}_{TSLS}^j + \varphi_j^V S_j^V}{\varphi_j^j - \varphi_j^B S_j^B - \tau p \hat{\beta}_{TSLS}^j},$$

where $\hat{\beta}_{TSLS}^j$ is the TSLS estimate of the effect of attending program j on test scores (in standard

deviation units), p is the predicted change in discounted lifetime earnings with a one standard

deviation test score increase, and τ represents the marginal tax rate. The first numerator term is

thus the present discounted value of the net-of-tax private earnings gains from participation for

the marginal program j attendee. The second is then the private transfer from program

substitution – the product of the marginal private program cost, φ_j^V, and the likelihood of

switching from a private center to program j, S_j^V. The denominator subtracts the marginal fiscal

savings from substitution across public programs, $\varphi_j^B S_j^B$, and the marginal discounted value of

future tax revenues, $\tau p \hat{\beta}_{TSLS}^j$, from marginal government outlays for program j, φ_j^j. It thus

captures the predicted cost to government of the marginal child's attendance, net of the fiscal

externalities from public program substitution and the additional tax revenue.[37]

The MVPF estimates in Table 9 Panel B are noisier than those for the benefit-to-cost

ratios but show a similar pattern to the estimates in Panel A. Estimates of the MVPF are larger

for universal pre-K and when program outlays are set equal to per-pupil Head Start spending.

Indeed, in that case, the MVPF of universal pre-K is 4.27. This estimate is much higher than that

seen for a number of redistributive policies (Hendren, 2016), or even for Head Start (Kline and

---

[37] The benefit-to-cost ratios presented in Table 9 Panel A can be represented by this notation. The first, gross-cost estimate is given by $p\hat{\beta}_{TSLS}^j / \varphi_j^j$. The second, net-cost estimate is given by $p\hat{\beta}_{TSLS}^j / \left(\varphi_j^j - \varphi_j^B S_j^B - \varphi_j^V S_j^V\right)$.

Walters, 2016). However, I have more confidence in the conclusion that the MVPF of universal pre-K truly differs from that of targeted pre-K under the assumption of higher (K-12) costs.

But comparing the MVPFs of universal and targeted programs would only be truly meaningful for a welfare analysis if they served the same populations. Put differently, if these programs both only served poor children, we could then draw the further conclusion that universal pre-K is the more desirable policy. While most of the true beneficiaries of universal pre-K appear to be poor children, I cannot rule out that higher-income children gain and so cannot draw this conclusion. Estimates of the MVPF for universal pre-K should therefore be helpful reference points for future work where monetary returns of universal pre-K attendance are measured in other data or perhaps more directly. By the same reasoning, however, my findings do not provide a strong case for targeted *state-funded* pre-K programs: the MVPF for targeted, state-funded pre-K programs is significantly lower than that for the federally-funded Head Start program (Kline and Walters, 2016), which serves essentially the same income group.

VIII.   Conclusion

This paper has presented new estimates of the impacts of preschool education in the U.S., and by harnessing the benefits of observing children from across the country in an underused longitudinal survey, moved the literature forward in meaningful ways. First, I have presented comparable estimates of the immediate cognitive test score impacts of both universal and targeted pre-K, based not only on the same data but also the same research design. I have found evidence that the benefits from attending state-funded pre-K in a state with a universal program exceed those from attending state-funded pre-K in a state with a targeted (means-tested) program – not just overall but especially for the poor children that both types of programs serve. I also attempted to rule out alternative explanations – beyond pre-K access – for the differences in pre-

K attendance impacts across these two groups of states. Throughout, my empirical approach addressed limitations of prior pre-K evaluations exploiting age-eligibility rules for identification. Most notably, survey data from the ECLS-B supported the use of a comparison group, estimation of valid TOT effects, and incorporation of new tests of internal validity.

The constellation of evidence is consistent with universal pre-K delivering greater benefits to the population it serves, relative to the costs, than targeted pre-K. In other words, there may be an efficiency-related justification for choosing a universal program over a targeted one, at least in the context of state-funded pre-K. The fact that universal pre-K delivers short-term benefits similar to universal kindergarten suggests that it looks considerably more like public education than targeted pre-K. Public education being perhaps the one investment in children that has the most political support in U.S., political economy considerations may be important for understanding universal pre-K's relative success.

This study is of course not without limitations. Though short-term cognitive test score gains from educational intervention predict impacts on adult outcomes like earnings (Chetty et al., 2011) – an idea that I use in the cost-benefit analysis – the findings in this paper only pertain to short-term cognitive effects for one birth cohort. Estimating the impacts of universal and targeted pre-K on longer-term outcomes, once the programs are mature enough to do so, would resolve uncertainty over whether the test score impacts documented here truly manifest in better outcomes over the longer term. Even looking at short-term outcome for more than one cohort would be helpful, if only because the present estimates are noisier than would be ideal. To my knowledge, however, the ECLS-B is the only dataset currently available in the U.S. to carry out this analysis.

In addition, while the findings of this paper provide a concrete example of a public

program where universal eligibility may raise cost efficacy, the implications might not extend to situations where public goods are not directly provided by the government, or even beyond early education. Given the idiosyncratic features of policy debates over access, it is important to study the impacts of eligibility rules on the output and productivity of other public programs directly.

References

Aizer, Anna and Flavio Cunha. 2012. "The Production of Human Capital: Endowments, Investments and Fertility." NBER Working Paper 18429.

Almond, Douglas, Janet Currie, and Valentina Duque. 2017. "Childhood Circumstances and Adult Outcomes: Act II." NBER Working Paper 23017.

Anderson, Patricia, Kristin Butcher, Elizabeth U. Cascio, and Diane Schanzenbach. 2011. "Is Being in School Better? The Impact of School on Children's BMI When Starting Age is Endogenous." *Journal of Health Economics* 30(5): 977-986.

Araujo, M. Caridad, Pedro Carneiro, Yyannu Cruz-Aguayo, and Norbert Schady. 2016. "Teacher Quality and Learning Outcomes in Kindergarten." *Quarterly Journal of Economics* 131(3): 1415-1453.

Araujo, M. Caridad, Marta Dormal, and Norbert Schady. Forthcoming. "Child Care Quality and Child Development." *Journal of Human Resources* 54(3).

Barnett, W. Steven, Jason T. Hustedt, Laura E. Hawkinson, and Kenneth B. Robin. 2006. *The State of Preschool 2006*. New Brunswick, NJ: The National Institute for Early Education Research.

Barnett, W. Steven, Allison H. Friedman-Krauss, G.G. Weisenfeld, Michelle Horowitz, Richard Kasmin, James H. Squires. 2017. *The State of Preschool 2016*. New Brunswick, NJ: The National Institute for Early Education Research.

Barr, Andrew and Chloe R. Gibbs. 2018. "Breaking the Cycle? Intergenerational Effects of an Anti-Poverty Program in Early Childhood." Mimeo.

Bitler, Marianne P., Hilary W. Hoynes, and Thurston Domina. 2014. "Experimental Evidence on Distributional Effects of Head Start." NBER Working Paper 20434.

Carneiro, Pedro and Rita Ginja. 2014. "Long-Term Impacts of Compensatory Preschool on Health and Behavior: Evidence from Head Start." *American Economic Journal: Economic Policy* 6(4): 135-173.

Cascio, Elizabeth U. 2015. "The Promises and Pitfalls of Universal Early Education." *IZA World of Labor* 116.

Cascio, Elizabeth U. and Diane Whitmore Schanzenbach. 2013. "The Impacts of Expanding Access to High-Quality Preschool Education." *Brookings Papers on Economic Activity*, Fall 2013, 127-178.

Chetty, Raj, John N. 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(4): 1593-1660.

Currie, Janet and Duncan Thomas. 1995. "Does Head Start Make a Difference?" *American Economic Review* 85(3): 341-364.

Deming, David. 2009. "Early Childhood Intervention and Life-Cycle Skill Development: Evidence from Head Start." *American Economic Journal: Applied Economics* 1(3): 111-134.

Dickert-Conlin, Stacy and Todd Elder. 2010. "Suburban Legend: School Cutoff Dates and the Timing of Births." *Economics of Education Review* 29(5): 826-841.

Elango, Sneha, Jorge Luis Garcia, James J. Heckman, and Andres Hojman. 2016. "Early Childhood Education." In Economics of Means-Testing Transfer Programs in the United States, Volume 2" (Ed. Robert Moffitt), pp. 235-297.

Elder, Todd and Darren Lubotsky. 2009. "Kindergarten Entrance Age and Children's Achievement: Impacts of State Policies, Family Background, and Peers." *Journal of Human Resources* 44(3): 641-683.

Engel, Mimi, Amy Claessens, and Maida A. Finch. 2013. "Teaching Students What They Already Know? The (Mis)Alignment Between Mathematics Instructional Content and Student Knowledge in Kindergarten." *Educational Evaluation and Policy Analysis* 35(2): 157-178.

Feller, Avi, Todd Grindal, Luke Miratrix, and Lindsay Page. 2016. "Compared to What? Variation in the Impacts of Early Childhood Education by Alternative Care Type." *The Annals of Applied Statistics* 10(3): 1245-1285.

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(1) (Advances), Article 46.

Fitzpatrick, Maria D. 2010. "Preschoolers Enrolled and Mothers at Work? The Effects of Universal Pre-Kindergarten." *Journal of Labor Economics* 28(1): 51-85.

Fitzpatrick, Maria D., David Grissmer, and Sarah Hastedt. 2011. "What a difference a day makes: Estimating daily learning gains during kindergarten and first grade using a natural experiment." *Economics of Education Review* 30: 269-279.

Flood, Sarah, Miriam King, Steven Ruggles, and J. Robert Warren. 2015. *Integrated Public Use Microdata Series, Current Population Survey: Version 4.0* [dataset]. Minneapolis, MN: University of Minnesota. http://doi.org/10.18128/D030.V4.0.

Garces, Eliana, Duncan Thomas, and Janet Currie. 2002. "Longer-Term Effects of Head Start." *American Economic Review* 92(4): 999-1012.

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(3): 533-558.

Heckman, James J., Seong Hyeok Moon, Rodrigo Pinto, Peter A. Savelyev, and Adam Yavitz. 2010. "The Rate of Return to the High/Scope Perry Preschool Program." *Journal of Public Economics* 94(1-2): 114-128.

Hendren, Nathaniel. 2016. "The Policy Elasticity." *Tax Policy and the Economy* 30(1): 51-89.

Johnson, Rucker C. and C. Kirabo Jackson. Forthcoming. "Reducing Inequality through Dynamic Complementarity: Evidence from Head Start and Public School Spending." *American Economic Journal: Economic Policy*.

Kline, Patrick and Christopher Walters. 2016. "Evaluating Public Programs with Close Substitutes: The Case of Head Start." *Quarterly Journal of Economics* 131(4): 1795-1848.

Ladd, Helen F., Clara G. Muschkin, and Kenneth A. Dodge. 2014. "From Birth to School: Early Childhood Initiatives and Third-Grade Outcomes in North Carolina." *Journal of Policy Analysis and Management* 33(1): 162-187.

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, Carol Bilbrey. 2013. *Evaluation of the Tennessee Voluntary Prekindergarten Program: Kindergarten and First Grade Follow-Up Results from the Randomized Control Design* (Research Report). Nashville, TN: Vanderbilt University, Peabody Research Institute.

Lipsey, Mark W., Christina Weiland, Hirokazu Yoshikawa, Sandra Jo Wilson, Kerry G. Hofer. 2015. "The Prekindergarten Age-Cutoff Regression-Discontinuity Design: Methodological Issues and Implications for Application." *Educational Evaluation and Policy Analysis*.

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(1): 159-208.

Puma, Michael, Stephen Bell, Ronna Cook, and Camilla Heid. 2010. *Head Start Impact Study Final Report*. U.S. Department of Health and Human Services, Administration for Children and Families. Washington, D.C.

Reardon, Sean. 2011. "The Widening Academic Achievment Gap between the Rich and the Poor: New Evidence and Possible Explanations." In Whither Opportunity: Rising Inequality, Schools, and Children's Life Changes (ed. Greg J. Duncan and Richard J. Murnane), pp. 91-116. New York: Russell Sage Foundation.

Ruggles, Steven, Katie Genadek, Ronald Goeken, Josiah Grover, and Matthew Sobek. 2015. *Integrated Public Use Microdata Series: Version 6.0* [dataset]. Minneapolis, MN: University of Minnesota. http://doi.org/10.18128/D010.V6.0.

Sabol, T.J., S.L. Soliday Hong, R.C. Pianta, and M.R. Burchinal. 2013. "Can Rating Pre-K Programs Predict Children's Learning?" *Science* (August): 845-846.

Schweinhart, Lawrence J., Jeanne Montie, Zongping Xiang, W. Steven Barnett, Clive R. Belfield, Milagros Nores. 2005. *Lifetime Effects: The High/Scope Perry Preschool Study Through Age 40*. Ypsilanti, MI: High/Scope Press.

Thompson, Owen. 2018. "Head Start's Long-Run Impact: Evidence from the Program's Introduction." *Journal of Human Resources* 53(4): 1100-1139.

Walters, Christopher. 2015. "Inputs in the Production of Early Childhood Human Capital: Evidence from Head Start." *American Economic Journal: Applied Economics* 7(4): 76-102.

Weiland, Christina and Hirokazu Yoshikawa. 2013. "Impacts of a prekindergarten program on children's mathematics, language, literacy, executive function, and emotional skills." *Child Development* 84(6): 2112-2130.

Wong, V.C., T.D. Cook, W.S. Barnett, and K. Jung. 2008. "An Effectiveness-Based Evaluation of Five State Pre-kindergarten Programs." *Journal of Policy Analysis and Management* 27(1): 122-154.

Figure 1. Trends in Public Preschool Enrollment Rates and Pre-K Funding: 1968-2015
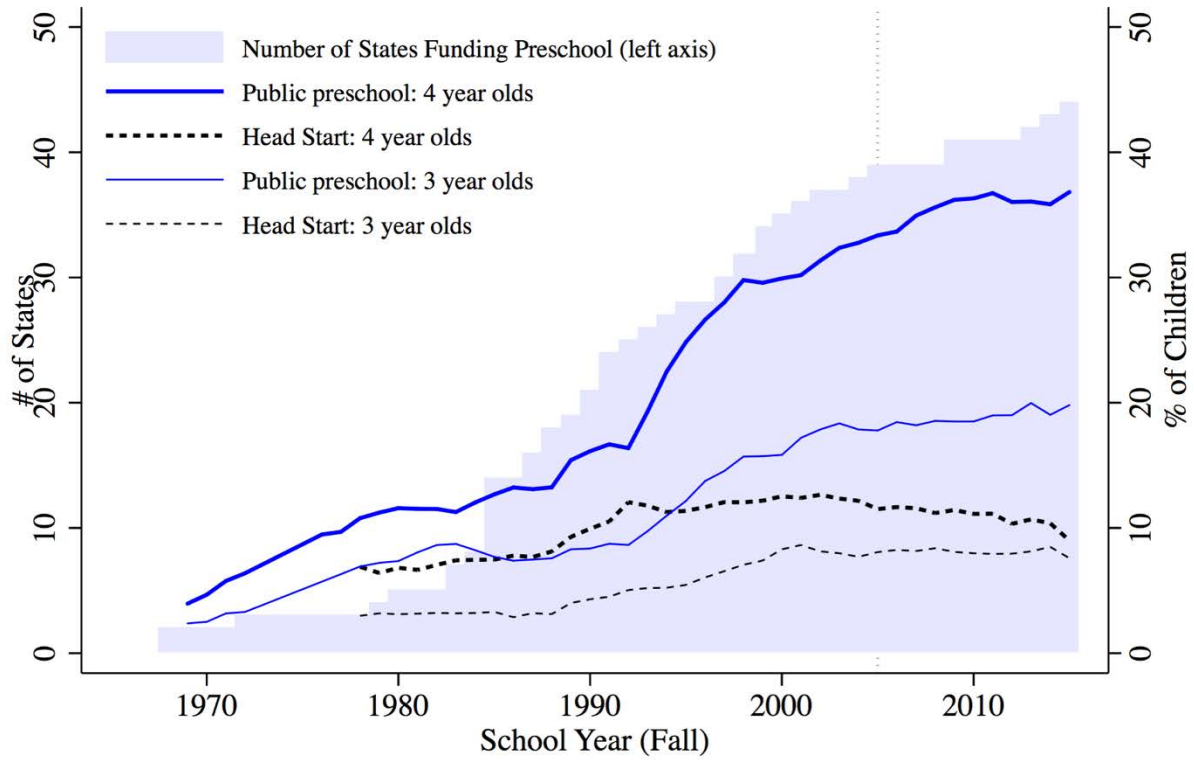
Notes: Data on public preschool enrollment rates by age are 3-year moving averages calculated from the 1968-2016 October Current Population Survey (CPS) School Enrollment supplements. Head Start enrollment rates divide Head Start enrollments reported by the Head Start Bureau by cohort size estimates based on annual (as of July 1) national age-specific population estimates from the Census Bureau. State funding dates were constructed from program narratives published by the National Institute for Early Education Research (Barnett et al., 2017).

Figure 2. Prekindergarten Access and Quality Standards by State: 2005-06
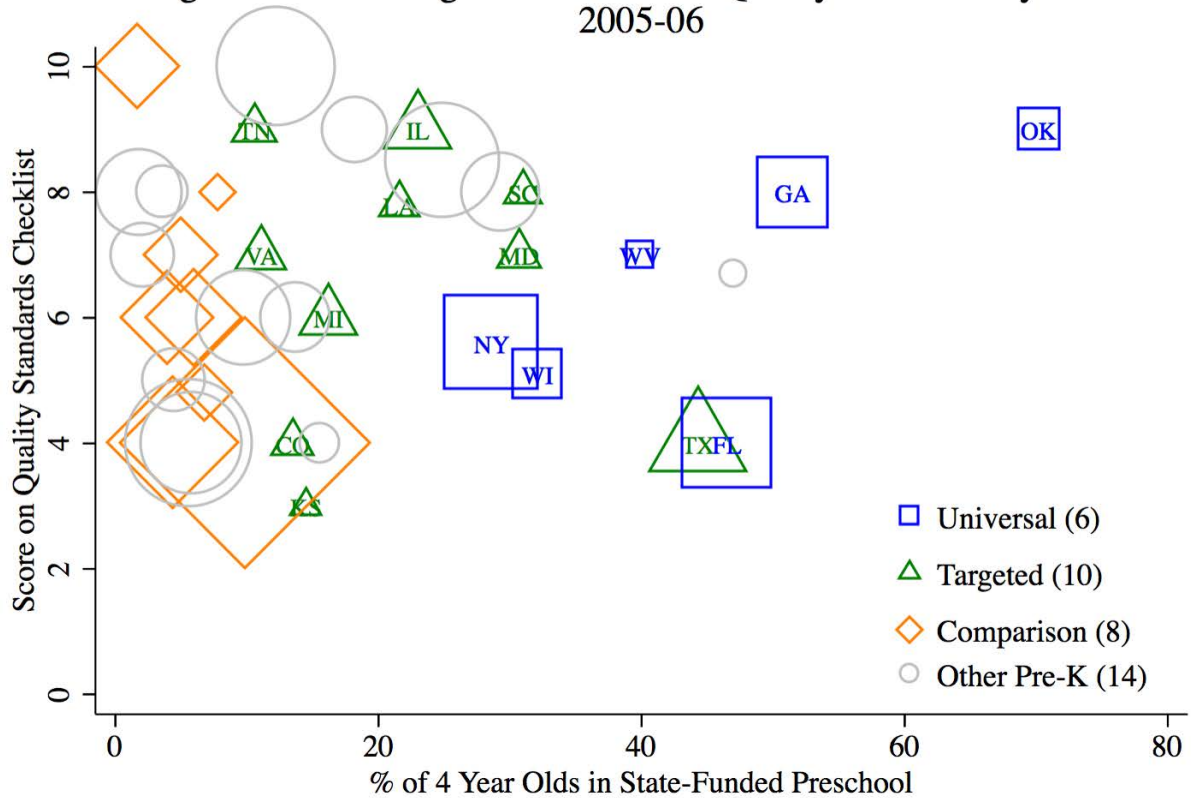
Notes: Data are from Barnett et al. (2006); dot sizes represent the size of the state's 4-year-old population. The quality standards checklist (vertical axis) has 10 points, one for each of 10 program standards: 1 point for comprehensive early learning standards, 4 points for teacher training and credentialing requirements (teacher has BA, specialized training in pre-K, assistant teacher has Child Development Associate or equivalent, at least 15 hours of in-service training annually), 2 points for staffing ratios (maximum class size no larger than 20, staff-child ratio 1:10 or better), 2 points for comprehensive services (vision, hearing, health, and one support service, at least one meal provided), and 1 point for a site visit requirement.

Figure 3. Pre-K Attendance and Test Scores
by Age and State Program Type

A. Pre-K Attendance
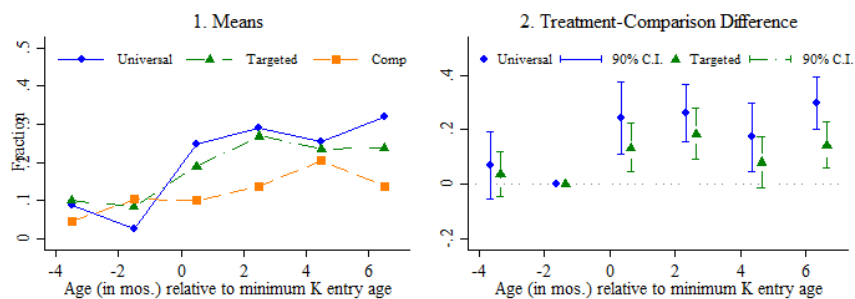
B. Preschool-Age Test Scores
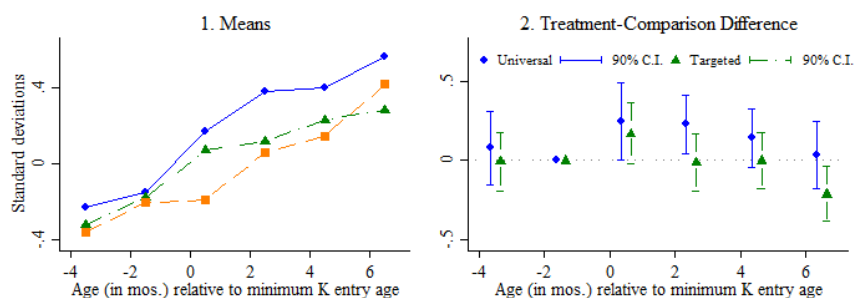
Notes: Data are from the ECLS-B. Estimation sample is restricted to respondents with non-missing values of key variables resident in one of the analysis states at wave 3, born with 4 months after and 8 months before that state's cutoff birthdate for kindergarten entry, and assessed during the 2005-06 school year. Panel A corresponds to pre-K attendance in 2005-06 (wave 3, when ECLS-B respondents were 4 years of age); Panel B corresponds to average standardized reading and math scores in 2005-06. Subpanel 1 of each panel plots average values of the dependent variable by age relative to the minimum age for kindergarten entry (2-month bins) for treated states with universal pre-K programs, treated states with targeted pre-K programs, and comparison states; see notes to Table 1 or Appendix Tables 1 and 2 for lists of states. The dots in subpanel 2 of each panel represent, separately for treatment states with universal programs and treatment states with targeted programs, the coefficients on interactions between a treatment dummy and a series of dummies for age relative to the minimum age for kindergarten entry (2-month bins) from a regression that allows for direct effects of each of these (sets of) variables in addition to month x year of assessment dummies and state fixed effects. The interaction with the dummy for missing eligibility by 1 to 2 months is omitted for identification. Capped vertical lines represent 90% confidence intervals, with standard errors clustered on state by month of birth. .

Figure 4. Eligibility Effects on Pre-K Attendance and Test Scores, by State Program Type and SES Quintile

A. Enrolled in Pre-K

B. Test Score: Preschool Age

Notes: Data are from the ECLS-B. Estimation sample is restricted to respondents with non-missing values of key variables resident in one of the analysis states at wave 3, born with 4 months after and 8 months before that state's cutoff birthdate for kindergarten entry, and assessed during the 2005-06 school year. Each dot in each panel represents an estimate of θ in model 1 restricting attention to children in states with universal or targeted programs (in both cases relative to the same group of comparison states) in the designated quintile of the ECLS-B index for socio-economic status. The SES index is measured contemporaneously with outcomes (in wave three, or 2005-06) and is derived from factor analysis on parental education, parental occupation, and family income. The underlying regression also includes indicators for month x year of assessment and the demographic and background characteristics listed in Table 1, Panel B. The capped vertical lines represent 90% confidence intervals, with standard errors clustered on state by month of birth.

Figure 5. Pre-K Eligibility and Test Scores by Age, State Program Type, and Poverty Status
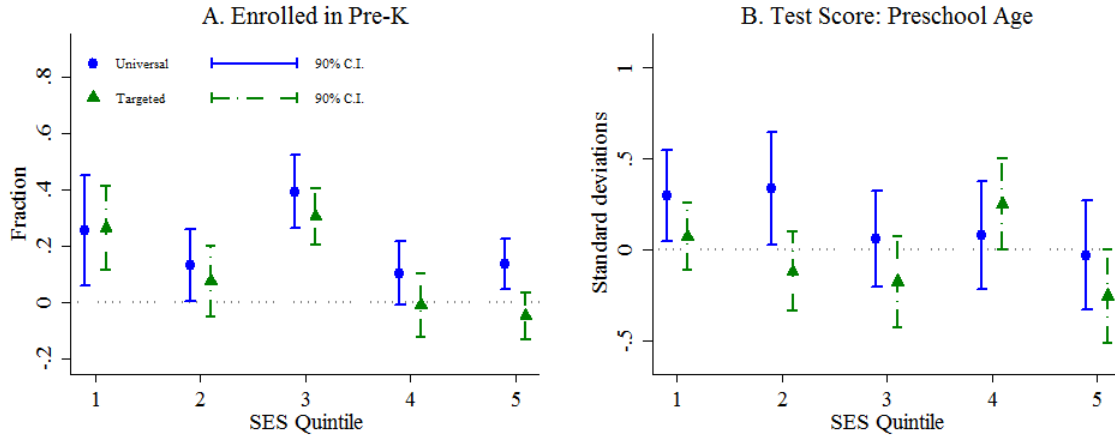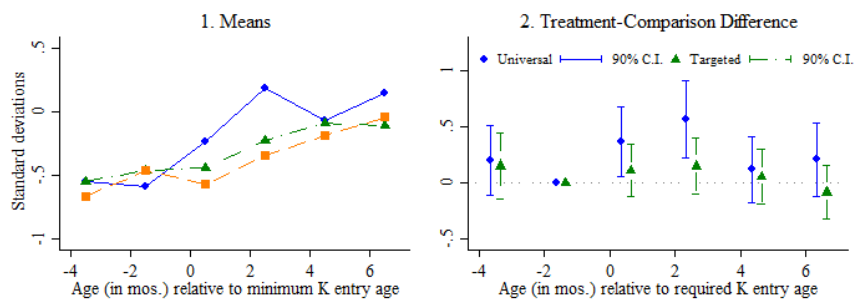
A. Low-Income (<=185% FPL)

B. Not Low-Income (> 185% FPL)
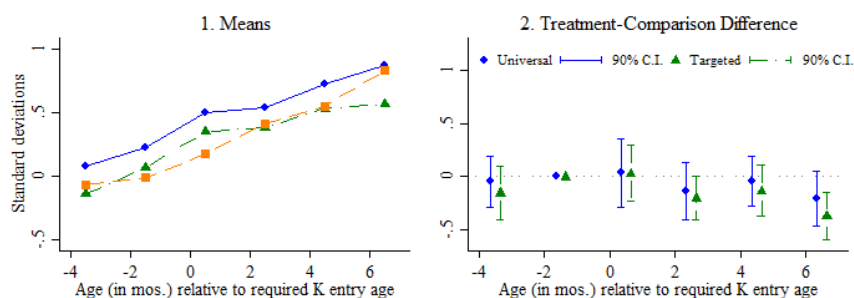
Notes: Data are from the ECLS-B. Estimation sample is restricted to respondents with non-missing values of key variables resident in one of the analysis states at wave 3, born with 4 months after and 8 months before that state's cutoff birthdate for kindergarten entry, and assessed during the 2005-06 school year. The dependent variable in each panel is average standardized reading and math scores during wave 3, when respondents were 4 years of age. Panel A corresponds to respondents who were eligible for free- or reduced-price lunch in 2005-06; Panel B corresponds to respondents who were not. Subpanel 1 of each panel plots the average standardized test score by age relative to the minimum age for kindergarten entry (2-month bins) for treated states with universal pre-K programs, treated states with targeted pre-K programs, and comparison states; see notes to Table 1 or Appendix Tables 1 and 2. The dots in subpanel 2 of each panel represent, separately for treatment states with universal programs and treatment states with targeted programs, the coefficients on interactions between a treatment dummy and a series of dummies for age relative to the minimum age for kindergarten entry (2-month bins) from a regression that allows for direct effects of each of these (sets of) variables in addition to month x year of assessment dummies and state fixed effects. The interaction with the dummy for missing eligibility by 1 to 2 months is omitted for identification. Capped vertical lines represent 90% confidence intervals, with standard errors clustered on state by month of birth.

Table 1. Descriptive Statistics and Balance Tests on Key Variables, by Program Type: Full Sample

	Universal		Targeted		Uni - Tar
	Ineligible Mean [sd]	DD Coef. (se)	Ineligible Mean [sd]	DD Coef. (se)	DDD Coef. (se)
	(1)	(2)	(3)	(4)	(5)
A. Treatment variable					
Pre-kindergarten[a]	0.052	0.211	0.091	0.114	0.097
		(0.047)		(0.033)	(0.048)
B. Background characteristics					
Age in months[a]	48.525	-0.099	48.235	-0.093	-0.006
	[3.160]	(0.060)	[3.316]	(0.051)	(0.055)
Female	0.463	0.078	0.521	-0.007	0.085
		(0.046)		(0.044)	(0.047)
Black non-Hispanic	0.206	0.042	0.201	0.026	0.015
		(0.033)		(0.029)	(0.035)
Hispanic	0.196	-0.034	0.255	-0.021	-0.014
		(0.039)		(0.033)	(0.036)
Low birth weight	0.074	0.022	0.079	0.015	0.007
		(0.014)		(0.012)	(0.014)
Maternal education ≤ HS[a]	0.445	0.022	0.468	-0.031	0.053
		(0.042)		(0.045)	(0.042)
Both biological parents in HH [a]	0.704	-0.052	0.692	0.045	-0.097
		(0.038)		(0.035)	(0.041)
Non-English at home[a]	0.152	-0.010	0.160	-0.008	-0.002
		(0.031)		(0.026)	(0.030)
Family income ≤ 185% FPL[a]	0.473	-0.092	0.451	-0.064	-0.028
		(0.043)		(0.043)	(0.045)
C. *p*-value: joint test for background chars.					
All background characteristics		0.01		0.17	0.26
Excluding poverty		0.32		0.40	0.27
Observations[b]	300	3,400	400	3,950	7,350

Notes: Odd-numbered columns give means for respondents in treatment states ineligible for pre-K in 2005-06 [standard deviations for non-binary variables]. Even-numbered columns give DD coefficients (standard errors) on the interaction between a dummy for being eligible for pre-K in 2005-06 and a dummy for residing in a treatment state (*elig* x *treat*) from a separate regression that also includes dummies for state of residence, month x year of assessment, and month age five relative to the state kindergarten entry cutoff birthdate in 2006-07. Treatment states are those with state-funded pre-K programs focused much more on 4-year-olds than 3-year-olds and statewide minimum age at pre-K entry cutoffs not in the middle of the month; treatment states with universal programs are FL, GA, NY, OK, WI, and WV, and treatment states with targeted programs are CO, IL, KS, LA, MI, MD, SC, TN, TX, and VA. Comparison states have statewide age at kindergarten entry regulations; some comparison states have relatively small pre-K programs (AL, CA, DE, MO, NM, OH, OR, WA), while others lack pre-K programs (AK, HI, ID, IN, MS, ND, RI, SD, UT). A child is deemed eligible for K in 2006-07 if he /she turned age 5 in time to start K in fall 2006, given his/her date of birth and the kindergarten entry age regulations in effect in 2006-07 reported by Barnett et al. (2007). Sample is limited to children who turn age 5 between 4 months after and 8 months before the cutoff date and who are assessed during the 2005-06 school year. Means and regressions are weighted by longitudinal sampling weights, and standard errors (in parentheses) are clustered on state x month of birth. [a] Measured at preschool age, or in 2005-06 (wave 3 interview). [b] rounded to the nearest 50, per IES guidelines.

Table 2. Impacts of State-funded Pre-K on Preschool-Age Test Scores

	Pre-K:	Test Score (Average Reading and Math):			
	First Stage	Ineligible Mean	RF (ITT)	IV (TOT)	OLS
	(1)	(2)	(3)	(4)	(5)
A. Universal (N=3,400)					
No additional controls	0.211	-0.183	0.144	0.682	-0.168
	(0.047)		(0.081)	(0.389)	(0.057)
With additional controls[a]	0.211		0.120	0.569	-0.069
	(0.047)		(0.065)	(0.315)	(0.048)
B. Targeted (N=3,950)					
No additional controls	0.114	-0.235	-0.009	-0.076	-0.232
	(0.033)		(0.069)	(0.594)	(0.049)
With additional controls[a]	0.118		-0.044	-0.373	-0.093
	(0.034)		(0.060)	(0.513)	(0.042)
C. p-value on Universal-Targeted Difference (N=7,350)					
No additional controls	0.045		0.071	0.185	0.227
With additional controls[a]	0.057		0.012	0.046	0.609

Notes: The first stage and RF (ITT) (reduced form (intent-to-treat)) columns in Panels A and B give coefficients on the interaction between a dummy for being eligible for kindergarten in 2006-07 (same as a dummy for being eligible for pre-K in 2005-06 in a treatment state) and a dummy for being in a treated state (*elig* x *treat*). (See Table 1 or Data Appendix for a definition of treatment and comparison states.) The IV (TOT) (instrumental variables (treatment on treated)) and OLS columns in Panels A and B give coefficients on pre-K attendance (*prek*), in the first case instrumenting with *elig* x *treat* using TSLS. All coefficients are from separate regressions that also include as controls dummies for state of residence, month x year of assessment, and month age five relative to the state kindergarten entry cutoff birthdate in 2006-07. Regressions are weighted by longitudinal sampling weights, and standard errors (in parentheses) are clustered on state x month of birth. [a] Additional controls include all demographic and background characteristis in Table 1 Panel B: age at assessment, dummies for female, Hispanic, black non-Hispanic, low birth weight, non-English at home, mom has high school degree or less, both biological parents in household, and low-income.

Table 3. Impacts of State-funded Pre-K on Preschool-Age Test Scores, by Poverty Status

| | Pre-K: First Stage (1) | Test Score (Average Reading and Math): | | | |
		Ineligible Mean (2)	RF (ITT) (3)	IV (TOT) (4)	OLS (5)
			A. Universal		
Low-income (N=1,550)	0.226 (0.062)	-0.572	0.263 (0.107)	1.160 (0.544)	-0.037 (0.055)
Not low-income (N=1,850)	0.189 (0.047)	0.165	0.007 (0.102)	0.038 (0.530)	-0.079 (0.077)
p-value on difference	0.550		0.118	0.191	0.635
			B. Targeted		
Low-income (N=1,750)	0.223 (0.051)	-0.497	-0.018 (0.102)	-0.082 (0.450)	-0.015 (0.051)
Not low-income (N=2,200)	0.037 (0.037)	-0.018	-0.066 (0.080)	no f.s.	-0.174 (0.079)
p-value on difference	0.002		0.732		0.107
		C. p-value on Universal-Targeted Difference			
Low-income	0.996		0.02	0.032	0.728
Not low-income	0.001		0.461	n.a.	0.255

Notes: A child is considered low income if his (preschool-age or 2005-06) family income is at or below 185% FPL, the threshold for eligibility for reduced-price lunch and the modal income-eligibility criterion for the targeted programs under study. The first stage and RF (ITT) (reduced form (intent-to-treat)) columns in Panels A and B give coefficients on the interaction between a dummy for being eligible for kindergarten in 2006-07 (same as a dummy for being eligible for pre-K in 2005-06 in a treatment state) and a dummy for being in a treated state (*elig* x *treat*). (See Table 1 or Data Appendix for a definition of treatment and comparison states.) The IV (TOT) (instrumental variables (treatment on treated)) and OLS columns in Panels A and B give coefficients on pre-K attendance (*prek*), in the first case instrumenting with *elig* x *treat* using TSLS. All coefficients are from separate regressions that also include as controls dummies for state of residence, month x year of assessment, and month age five relative to the state kindergarten entry cutoff birthdate in 2006-07, and all demographic and background characteristics listed in Table 1 Panel B except the low-income indicator. Regressions are weighted by longitudinal sampling weights, and standard errors (in parentheses) are clustered on state x month of birth. "no f.s." means "no first stage." "n.a." means "not applicable."

Table 4. Sensitivity of Estimated Effects of Pre-K to Choice of Outcome

	Pre-K: First Stage (1)	Test Scores		
		RF (ITT) (2)	IV (TOT) (3)	OLS (4)
A. Baseline				
Universal (N=3,400)	0.211	0.120	0.569	-0.069
	(0.047)	(0.065)	(0.315)	(0.048)
Targeted (N=3,950)	0.118	-0.044	-0.373	-0.093
	(0.034)	(0.060)	(0.513)	(0.042)
p-value on difference	0.057	0.012	0.046	0.432
B. Reading score only				
Universal (N=3,400)	0.211	0.135	0.639	-0.036
	(0.047)	(0.064)	(0.328)	(0.047)
Targeted (N=3,950)	0.118	-0.037	-0.314	-0.100
	(0.034)	(0.067)	(0.564)	(0.045)
p-value on difference	0.057	0.007	0.054	0.214
C. Math score only				
Universal (N=3,400)	0.211	0.106	0.500	-0.102
	(0.047)	(0.082)	(0.378)	(0.058)
Targeted (N=3,950)	0.118	-0.051	-0.432	-0.087
	(0.034)	(0.067)	(0.587)	(0.047)
p-value on difference	0.057	0.061	0.102	0.786
D. Mental score (age 2)[a]				
Universal (N=3,350)	0.211	0.005	0.023	-0.062
	(0.046)	(0.069)	(0.326)	(0.051)
Targeted (N=3,950)	0.115	0.034	0.299	-0.081
	(0.034)	(0.068)	(0.610)	(0.051)
p-value on difference	0.049	0.613	0.556	0.643
E. = 1 if parent reports not ready for K (age 4)				
Universal (N=3,400)	0.207	-0.091	-0.440	0.003
	(0.047)	(0.046)	(0.231)	(0.025)
Targeted (N=3,950)	0.114	-0.036	-0.315	-0.021
	(0.033)	(0.036)	(0.343)	(0.023)
p-value on difference	0.053	0.261	0.712	0.327

Notes: The first stage and RF (ITT) (reduced form (intent-to-treat)) columns give coefficients on the interaction between a dummy for being eligible for kindergarten in 2006-07 (same as a dummy for being eligible for pre-K in 2005-06 in a treatment state) and a dummy for being in a treated state (*elig* x *treat*). (See Table 1 or Data Appendix for a definition of treatment and comparison states.) The IV (TOT) (instrumental variables (treatment on treated)) and OLS columns give coefficients on pre-K attendance (*prek*), in the first case instrumenting with *elig* x *treat* using TSLS. All coefficients are from separate regressions that also include as controls dummies for state of residence, month x year of assessment, and month age five relative to the state kindergarten entry cutoff birthdate in 2006-07, and all demographic and background characteristics listed in Table 1 Panel B. Regressions are weighted by longitudinal sampling weights, and standard errors (in parentheses) are clustered on state x month of birth. [a] Additional controls include age at assessment and dummies for month x year of assessment in wave 2 of the ECLS-B, and panel weights incorporate observation in wave 2.

Table 5. Sensitivity of Estimated Effects of Pre-K on Test Scores to Estimation Sample

	Pre-K: First Stage (1)	Test Scores RF (ITT) (2)	IV (TOT) (3)	OLS (4)
A. Baseline				
Universal (N=3,400)	0.211	0.120	0.569	-0.069
	(0.047)	(0.065)	(0.315)	(0.048)
Targeted (N=3,950)	0.118	-0.044	-0.373	-0.093
	(0.034)	(0.060)	(0.513)	(0.042)
p-value on difference	0.057	0.012	0.046	0.432
B. +/- 4 months from threshold[a]				
Universal (N=2,000)	0.226	0.203	0.900	-0.05
	(0.054)	(0.075)	(0.350)	(0.071)
Targeted (N=2,350)	0.129	0.042	0.321	-0.067
	(0.039)	(0.068)	(0.527)	(0.056)
p-value on difference	0.091	0.024	0.197	0.808
C. Expanded sample (+ middle of month cutoffs)[b]				
Universal (N=3,450)	0.212	0.127	0.600	-0.046
	(0.045)	(0.065)	(0.314)	(0.048)
Targeted (N=4,250)	0.118	-0.020	-0.170	-0.089
	(0.032)	(0.059)	(0.503)	(0.040)
p-value on difference	0.042	0.022	0.089	0.379
D. Restricted sample (- late cutoffs)[c]				
Universal (N=1,950)	0.289	0.128	0.442	-0.082
	(0.054)	(0.080)	(0.274)	(0.068)
Targeted (N=2,550)	0.191	-0.036	-0.188	-0.120
	(0.036)	(0.073)	(0.375)	(0.054)
p-value on difference	0.078	0.016	0.037	0.622

Notes: The first stage and RF (ITT) (reduced form (intent-to-treat)) columns give coefficients on the interaction between a dummy for being eligible for kindergarten in 2006-07 (same as a dummy for being eligible for pre-K in 2005-06 in a treatment state) and a dummy for being in a treated state (*elig* x *treat*). (See Table 1 or Data Appendix for a definition of treatment and comparison states in the baseline sample.) The IV (TOT) (instrumental variables (treatment on treated)) and OLS columns give coefficients on pre-K attendance (*prek*), in the first case instrumenting with *elig* x *treat* using TSLS. All coefficients are from separate regressions that also include as controls dummies for state of residence, month x year of assessment, and month age five relative to the state kindergarten entry cutoff birthdate in 2006-07, and all demographic and background characteristics listed in Table 1 Panel B. Regressions are weighted by longitudinal sampling weights, and standard errors (in parentheses) are clustered on state x month of birth. [a] Sample further limited to respondents with birthdays within 4 months of the cutoff birthdate for kindergarten entry in their wave 3 state of residence. [b] Sample expanded to include respondents residing in treatment states with middle-of-month cutoffs (targeted: AR, NC; universal: ME) not born in the cutoff birthdate month. [c] Sample eliminates treatment and comparison states with cutoff birthdates after October 1 (universal: NY, targeted: MI, comparison: CA, HI).

Table 6. Impacts on Alternative Care Arrangements, by State Program Type: Low-Income Children

	Pre-K: First Stage (1)	Head Start (2)	Other center-based care (3)	Any formal (2+3) (4)	Informal non-parental care (5)	Parental care (6)	Any informal (5+6) (7)
				Alternatives:			
				A. Ineligible Means			
Universal	0.08	0.22	0.21	0.43	0.18	0.31	0.49
Targeted	0.11	0.21	0.10	0.32	0.18	0.40	0.57
				B. Reduced form (coef (se) on elig x treat$_s$)			
Universal (N=1,550)	0.226	-0.079	-0.022	-0.102	-0.033	-0.092	-0.125
	(0.062)	(0.070)	(0.058)	(0.088)	(0.047)	(0.072)	(0.068)
Targeted (N=1,750)	0.223	-0.082	0.016	-0.066	-0.038	-0.120	-0.158
	(0.051)	(0.063)	(0.051)	(0.069)	(0.045)	(0.059)	(0.060)
p-value on difference	0.966	0.968	0.483	0.711	0.924	0.700	0.647
				C. Instrumental Variables (coef (se) on prek$_{is}$)			
Universal (N=1,550)	n.a.	-0.349	-0.099	-0.448	-0.146	-0.406	-0.552
		(0.274)	(0.242)	(0.317)	(0.203)	(0.327)	(0.317)
Targeted (N=1,750)	n.a.	-0.368	0.073	-0.295	-0.168	-0.537	-0.705
		(0.244)	(0.227)	(0.273)	(0.196)	(0.268)	(0.273)
p-value on difference	n.a.	0.948	0.459	0.660	0.914	0.704	0.660

Notes: A child is considered low income if his (preschool-age or 2005-06) family income is at or below 185% FPL, the threshold for eligibility for reduced-price lunch and the modal income-eligibility criterion for the targeted programs under study. Reduced-form coefficients in Panel A are on the interaction between a dummy for being eligible for kindergarten in 2006-07 (same as a dummy for being eligible for pre-K in 2005-06 in a treatment state) and a dummy for being in a treated state (*elig* x *treat*). (See Table 1 or Data Appendix for a definition of treatment and comparison states.) The instrumental variables coefficients in Panel B are on pre-K attendance (*prek*), estimated by instrumenting with *elig* x *treat* using TSLS. All coefficients are from separate regressions that also include as controls dummies for state of residence, month x year of assessment, and month age five relative to the state kindergarten entry cutoff birthdate in 2006-07, and all demographic and background characteristics listed in Table 1 Panel B except the low-income indicator. Regressions are weighted by longitudinal sampling weights, and standard errors (in parentheses) are clustered on state x month of birth. "n.a." means "not applicable

Table 7. Robustness of the Universal-Targeted Difference in Age 4 Test Estimates to Regression-Adjusting for Other Sources of State Heterogeneity in Impacts of State-Funded Pre-K

	Full Sample			Low-Income		
	Test Scores		Characteristic:	Test Scores		Characteristic:
	RF (ITT)	IV (TOT)	RF	RF (ITT)	IV (TOT)	RF
	(1)	(2)	(3)	(4)	(5)	(6)
Baseline	0.164	0.942	n.a.	0.281	1.241	n.a.
	(0.065)	(0.473)		(0.120)	(0.580)	
Allowing for heterogeneity by:						
			A. Heterogeneity by Class Size Requirements			
# Class size standards	0.135	1.154	0.125	0.231	1.196	0.214
	(0.071)	(0.521)	(0.067)	(0.114)	(0.590)	(0.108)
			B. Heterogeneity by Individual Background Characteristics			
Female	0.162	1.090	0.088	0.279	1.235	0.062
	(0.065)	(0.498)	(0.063)	(0.120)	(0.576)	(0.095)
Black non-Hispanic	0.163	0.994	0.121	0.282	1.248	0.200
	(0.065)	(0.485)	(0.091)	(0.121)	(0.576)	(0.114)
Hispanic	0.154	0.907	0.023	0.267	1.190	-0.058
	(0.066)	(0.530)	(0.086)	(0.120)	(0.600)	(0.113)
Low birth weight	0.164	0.952	0.022	0.281	1.243	0.102
	(0.065)	(0.479)	(0.083)	(0.120)	(0.578)	(0.097)
Maternal education ≤ HS	0.160	1.313	0.172	0.277	1.233	0.200
	(0.065)	(0.662)	(0.074)	(0.121)	(0.567)	(0.103)
Both biological parents in HH	0.157	0.674	-0.143	0.276	1.228	-0.059
	(0.065)	(0.449)	(0.072)	(0.119)	(0.579)	(0.096)
Non-English at home	0.165	1.010	0.119	0.282	1.255	0.111
	(0.065)	(0.493)	(0.089)	(0.120)	(0.577)	(0.114)
Family income ≤ 185% FPL	0.168	1.314	0.116	n.a.	n.a.	n.a.
	(0.067)	(0.691)	(0.072)			
All characteristics	0.143	1.104	n.s.	0.256	1.198	n.s.
	(0.066)	(0.703)		(0.119)	(0.598)	

Notes: A child is considered low income if his (preschool-age or 2005-06) family income is at or below 185% FPL, the threshold for eligibility for reduced-price lunch and the modal income-eligibility criterion for the targeted programs under study. RF (ITT) columns give the difference in reduced-form (intent-to-treat) coefficients on *elig* x *treat* between universal and targeted states. IV (TOT) columns give the difference in instrumental variables (treatment-on-treated) coefficients on *prek* (instrumented with *elig* x *treat* using TSLS) between universal and targeted states. The RF column shows the coefficient on *elig* x *treat* x characteristic from the same model that produced the RF (ITT) coefficients. All regressions also include as controls dummies for state of residence, month x year of assessment, and month age five relative to the state kindergarten entry cutoff birthdate in 2006-07, and all demographic and background characteristics listed in Table 1 Panel B (except the low-income indicator in columns 4-6), with separate control coefficients for the universal and targeted estimation samples. Regressions are weighted by longitudinal sampling weights, and standard errors (in parentheses) are clustered on state x month of birth. "n.a." means "not applicable." "n.s." means "not shown."

Table 8. Impacts of Universal Public Kindergarten on Kindergarten-Age Test Scores, Overall and by Poverty Status

	Public K:	Test Score (Average Reading and Math):		
	First Stage	RF (ITT)	IV (TOT)	OLS
	(1)	(2)	(3)	(4)
		A. Full Sample (N=2,400)		
All states	0.657	0.376	0.573	0.250
	(0.036)	(0.084)	(0.132)	(0.064)
		B. Low-Income (N=1,050)		
All states	0.778	0.440	0.566	0.458
	(0.050)	(0.144)	(0.182)	(0.107)
		C. Not Low-Income (N=1,300)		
All states	0.557	0.332	0.597	0.110
	(0.049)	(0.095)	(0.180)	(0.073)

Notes: A child is considered low income if his (preschool-age or 2005-06) family income is at or below 185% FPL, the threshold for eligibility for reduced-price lunch. The first stage and RF (ITT) (reduced form (intent-to-treat)) columns in Panels A and B give coefficients on a dummy for being eligible for kindergarten in 2006-07 (*elig*). The IV (TOT) (instrumental variables (treatment on treated)) and OLS columns in Panels A and B give coefficients on public K attendance, in the first case instrumenting with *elig* using TSLS. All coefficients are from separate regressions that also include as controls dummies for state of (wave 4) residence and month x year of (wave 4) assessment, a linear term in month age five relative to the state kindergarten entry cutoff birthdate in 2006-07, an interaction between that linear term and *elig*, and all demographic and background characteristics listed in Table 1 Panel B (except the low-income dummy in Panels B and C). Regressions are weighted by longitudinal sampling weights, and standard errors (in parentheses) are clustered on state x month of birth. Sample is limited to the subset of the original estimation sample born within 4 months after and 4 months before the minimum kindergarten entry age in their wave 4 state of residence.

Table 9. Cost-Benefit Analysis Under Alternative Assumptions

	Universal Public K (1)	Universal Pre-K (2)	Targeted Pre-K (3)	p: Universal = Targeted (4)
A. Benefit-to-cost				
1. Program outlays = per-pupil K-12 spending				
Cost=program outlays	1.98	1.39	-0.18	0.16
	(0.45)	(0.77)	(1.03)	
	[0.03]	[0.61]	[0.25]	
Cost=net program outlays	2.96	1.85	-0.24	0.15
	(0.68)	(1.02)	(1.32)	
	[0.00]	[0.41]	[0.35]	
2. Program outlays = p.p. Head Start spending				
Cost=program outlays	n.a.	2.17	-0.27	0.14
		(1.20)	(1.50)	
		[0.33]	[0.40]	
Cost=net program outlays	n.a.	3.52	-0.39	0.12
		(1.94)	(2.19)	
		[0.20]	[0.52]	
B. Marginal Value of Public Funds (MVPF)				
1. Program outlays = per-pupil K-12 spending	3.94	1.96	-0.23	0.14
	(1.63)	(1.39)	(0.99)	
	[0.07]	[0.49]	[0.21]	
2. Program outlays = p.p. Head Start spending	n.a.	4.27	-0.39	0.25
		(4.19)	(1.57)	
		[0.43]	[0.38]	

Notes: Calculations of benefit-cost ratios and the MVPF consider two scenarios: one where the per-pupil program outlays are equal to per-pupil K-12 spending (an upper bound for pre-K programs) and another where per-pupil program outlays are equal to per-pupil Head Start spending. For the benefit-cost ratio calculation, benefits (numerator) are calculated by multiplying the preferred IV estimates of the per-pupil attendance impact on test scores (with additional controls, poor subsample in the case of targeted pre-K) by the predicted effect of a one-standard deviation increase in early life test scores on lifetime earnings; net costs (denomintor) subtract fiscal savings from substitution from Head Start and from other private center-based care. For the MVPF, the beneficiary's marginal willingess to pay (numerator) sums predicted earnings impacts net of taxes and the transfer to parents from substitution from other private center-based care; net government costs (denominator) subtract the fiscal externalities from substitution from Head Start and predicted tax revenues from per-pupil program outlays. Throughout, I adopt Kline and Walters (2016) conservative assumption that a one standard deviation increase in test scores yields a 10% increase in earnings; that targeted pre-K attendees have lifetime earnings approximately 80% of those for the average child; and that the PDV of lifetime earnings at age 12 is on average $522,000 (in 2010 dollars, from Chetty et al., 2011). I also assume a discount rate of 3% to calculate PDVs of lifetime earnings at age 5 (column 1) and age 4 (columns 2-3), a tax rate of 20%, and a per-pupil annual price of other, private center-based care equal to approximately $5000 (based on Laughlin (2013) calculations from the Survey of Income and Program Participation). Per-pupil K-12 and Head Start spending is from Barnett et al. (2006) and corresponds to the 2005-06 for the relevant states. Standard errors (in parentheses) are calculated using the delta method; p-values [in brackets] correspond to a test of whether the ratio equal to one.

Appendix A: Data

A.   Defining "Treatment" States

My analysis focuses on state programs that served 4-year-olds nearly exclusively, according to statistics and program narratives published by the National Institute for Early Education Research (NIEER) (Barnett et al., 2006).[38] I define such programs as those for which the difference in NIEER-reported state pre-K enrollment rates between 4- and 3-year-olds in 2005-06 was at least 8 percentage points. This definition is necessitated by my research design and the statistical power afforded by the data. In particular, it must be the case that there is a noticeable change in the likelihood of pre-K attendance in 2005-06 for children with birthdates near school entry cutoff dates. Even if a program serves only 4-year-olds, it will be hard to detect an effect on attendance if the overall 4-year-old enrollment rate in the program is low. Likewise, even if a program serves a high share of 4-year-olds, an attendance effect will be difficult to detect if it also serves a high share of 3-year-olds. Twenty-three of the 38 states with state-funded programs in 2005-06 met the criterion laid forth above, Illinois being the last (Appendix Table 1).

The data and research design also necessitate that I focus on states for which there were state-established dates by which the youngest enrollees were to have been 4-years-old; locally established cutoff dates are not provided available. In addition, the ECLS-B does not provide information on day of birth within a given month. To minimize misclassification in assignment of the kindergarten eligibility indicator, I therefore focus on states with cutoff birthdates at the beginning or end of the month in my main analysis. Four states (Vermont, Kentucky, Connecticut, and New Jersey) must be excluded due to local determination of their pre-K entry cutoff dates; three (Arkansas, Maine, and North Carolina) are excluded for having cutoff birthdates in the middle of the month. The latter states are included in a specification check in Table 5 (full sample) and Appendix Table 8 (poor subsample).

Thus, sixteen states meet all of the above criteria for 2005-06.[39] Appendix Table 1 lists these treatment states by whether they are universal or targeted, according to the NIEER narratives, in descending order according to their enrollment rate. In addition to Georgia and Oklahoma, the six universal states under consideration include Florida, New York, West Virginia, and Wisconsin. There are ten targeted programs under consideration. Five of these (those in Colorado, Louisiana, Maryland, Tennessee and Texas) use eligibility for free- or reduced-price lunch as their income eligibility requirement, though that requirement need not apply to all enrollees. Two of these (Kansas and Michigan) use different income eligibility requirements. The remaining three states (Illinois, South Carolina, and Virginia) have no explicit income requirements, but rather risk factor requirements that arguably correlate strongly with income.[40]

---

[38] I also use data from Barnett et al. (2007) on kindergarten entry cutoff birthdates for fall 2006.

[39] Unfortunately, Barnett et al. (2006) does not report on Washington, D.C., so I cannot include ECLS-B observations from D.C. in the analysis.

[40] Such risk factors include (but are not limited to) low parental education, single parenthood, English language learner (ELL) status, homelessness, placement in foster care, and developmental delays, depending on the state.

I use 17 comparison states (Appendix Table 2). Eight of these states (Alabama, California, Delaware, Missouri, New Mexico, Ohio, Oregon, and Washington) had pre-K programs in 2005-06 that were too small or not different enough in terms of enrollment of 3- and 4-year-olds to be included in the treatment group.[41] The remaining 9 comparison states (Alaska, Hawaii, Idaho, Indiana, Mississippi, North Dakota, Rhode Island, South Dakota, and Utah) did not have pre-K programs in 2005-06.

B.   ECLS-B Estimation Sample

The empirical approach taken in this paper is made possible by detailed survey data from the Birth Cohort sample of the Early Childhood Longitudinal Study (ECLS-B). The ECLS-B is a longitudinal survey of a stratified random sample of children born in the United States in 2001.[42] ECLS-B respondents were assessed and their parents and caregivers interviewed at roughly 9 months of age (wave 1), 2 years of age/toddler age (wave 2), 4 years of age/preschool age (wave 3), and kindergarten age (waves 4 and 5). My estimation sample consists of all children with non-missing preschool-age cognitive assessments and demographic and background characteristics residing at preschool age in one of the 16 treatment states or 17 comparison states, 5 years old between 8 months before and 4 months after their state's kindergarten entry cutoff, and assessed during the 2005-06 school year – a total of 5,100 observations.[43] Reported sample sizes are rounded to the nearest 50, per IES rules to protect confidentiality of ECLS-B respondents.

Most pertinent for this study are the data from wave 3; this wave includes test scores on children who were of preschool age, but may or may not have been actually enrolled in (or age-eligible to enroll in) state-funded pre-K. More specifically, given the fall kindergarten entry cutoffs in Appendix Table 1, the 2001 birth cohort can be split into two school entry cohorts in the wave 3 data – children eligible to enter kindergarten in fall 2006 (and pre-K in fall 2005, if relevant) and children eligible to enter kindergarten in fall 2007 (and pre-K in fall 2006, if relevant). Children in the estimation sample were then tested starting in September 2005 – when any exposure to pre-K would have been limited – through June 2006 – when a child enrolled would have had a full school year of exposure.

C.   Key Variables

The main outcomes of interest are cognitive test scores in early math and reading. The preschool cognitive assessment was designed to test both for developmental (age-based) milestones and for

---

[41] In a previous version of this paper (Cascio, 2017), the comparison group also included five states with pre-K programs (Arizona, Connecticut, Kentucky, Minnesota, and Nevada) with local rules regarding age eligibility for pre-K but statewide cutoff dates for kindergarten entry that were not in the middle of the month. Because many localities may choose a common cutoff date, like September 1, inclusion of these states in the comparison group weakens the first stage, so they were eliminated from the comparison group in the present version of the paper.
[42] The ECLS-B contains oversamples of some demographic groups (Chinese and other Asians, Pacific Islanders, Native Americans, and Alaskan Natives), twins, and low and very low birth weight children. I apply sampling weights to make the estimates population representative.
[43] Since the same comparison group is used to estimate the impact of targeted and universal programs, the number of observations may appear to be larger when summing across regression-specific sample sizes.

knowledge and skills considered important for school readiness and early school success.[44] For my main estimates, I work with reading and math scale scores from this assessment that I normalized to mean zero and variance one in the subsample of comparison states. The main test score outcome is the unweighted average of these standardized scores for reading and math. For estimation of the impacts of kindergarten attendance (Table 8), I consider outcomes from the kindergarten-age (wave 4) cognitive assessment standardized in same fashion. According to the ECLS-B guidance on the scores, "the majority of items in the kindergarten 2006/2007 battery come from the ECLS-K, although several preschool items were included in order to link the data waves within the ECLS-B cohort."

The ECLS-B also contains rich family background information on respondents.[45] In addition to basic demographics (age at assessment and indicators for sex (female) and race (non-Hispanic black and Hispanic)), I construct indicators for low-income (family income at or below 185% of the federal poverty line (FPL)), for low birth weight (birth weight < 2,500 grams), for low maternal education (at or below a high school degree), for a language other than English being spoken in the home, and for the presence of both biological parents in the household. I use the low-income indicator to stratify the analysis and as a control in the full-sample estimates and the remaining background and demographic characteristics as controls in my preferred model. The low-income indicator is ideal for stratification, since free or reduced-price lunch eligibility (family income ≤ 185% FPL) is the modal income eligibility criterion for the targeted states of interest.

The ECLS-B also assessed toddler's motor and mental development using the Bayley Short Form-Research Edition (based on the Bayley Scales of Infant Development, 2nd Edition). I standardize these measures analogously to the preschool-age test scores for the specification checks in Table 4 (full sample) and Appendix Table 4 (low-income subsample). I also include these pretest scores as additional controls in a specification check in Appendix Table 5. I focus on the mental scores but impacts on and controlling for the motor scores yield similar results.

The ECLS-B also provides detailed information on the care and education of respondents at preschool age. In particular, it provides (1) Parent reports of center characteristics for the center in which a child spends the most time; and (2) Provider reports of center characteristics for the type of non-parental care in which a child spends the most time. Source (1) is available for most children, whereas source (2) is missing for a decent share of children for which it should be available.

These data constraints guide my approach to calculating the dummy for state-funded pre-K attendance. I begin with a parent report of the child attending a free program or center.[46] But a concern with using this measure alone is that parents might classify a state-funded pre-K in a different way. For example, the fact that some states allow/encourage subcontracting with Head Start means some children who parents report to be in Head Start are actually enrolled in a state-

---

[44] The preschool-age assessment drew on the Peabody Picture Vocabulary Test (PPVT), the Preschool Comprehensive Test of Phonological and Print Processing (Pre-CTOPPP), the PreLAS® 2000, and the Test of Early Mathematics Ability-3 (TEMA-3), as well as the cognitive assessment given to the fall 1998 kindergarten cohort of the ECLS (ECLS-K).
[45] All time- or age-varying family background variables are measured in wave 3, or at preschool age.
[46] Most of these are classified as pre-K or preschool by the parent; only a small share are day care or nursery school.

funded pre-K. In cases where a parent does not report that a child is attending a free program or center, I therefore recode the pre-K indicator to one if the provider reports that a program is a public-school pre-K or a preschool (or other program) sponsored by state/local government or public schools. For example, the parent of a child attending pre-K via Head Start might report that the child is in Head Start, but not another center type. The second step recodes these as state-funded pre-K when Head Start is the primary source of non-parental care. Reassuringly, the second step changes the pre-K attendance dummy much more often for low-income children. Provider reports with which to make the second correction are sometimes missing. Parents also may not accurately report on their child's enrollment.

Appendix B: Cost-Benefit Analysis

The cost-benefit analysis in Table 9, discussed in Section VII, relies on a number of assumptions. The table below outlines the choices of key parameter values.

Estimate/parameter	Value	Source/Justification
$\hat{\beta}_{TSLS}^{U}$	0.569 (0.315)	Table 2 Panel A, additional controls, column 4
$\hat{\beta}_{TSLS}^{T}$	-0.082 (0.450)	Table 3, Panel B, low-income column 4
$\hat{\beta}_{TSLS}^{K}$	0.573 (0.132)	Table 8, Panel A, column 3
S_{U}^{V}, S_{U}^{B}	-0.448, -0.09	Appendix Table 9, Panel C, columns 2 and 3
S_{T}^{V}, S_{T}^{B}	0, -0.368	Table 6, Panel C, columns 2 and 3
S_{K}^{V}, S_{K}^{B}	-0.378, -0.24	Appendix Table 10, Panel C, columns 2+5 and 4
φ_{U}^{U}	$11,875 (K-12), $7627 (Head Start)	Age 4 population-weighted averages of state figures in Barnett et al. (2006)
φ_{T}^{T}	$10,139 (K-12), $6974 (Head Start)	Age 4 population-weighted averages of state figures in Barnett et al. (2006)
φ_{K}^{K}	$11,000	Approximation
$\varphi_{U}^{V} = \varphi_{K}^{V}$	$4,966	Lower-bound estimate from Laughlin (2013)
$\varphi_{U}^{B}, \varphi_{T}^{B}, \varphi_{K}^{B}$	$7627, $6974, $7,300	Age 4 population-weighted averages of state per-pupil Head Start spending in Barnett et al. (2006)
p	$0.1 \times e_j$, where is the present discounted value of lifetime earnings at age 4 ($j=U,T$) or age 5 ($j=K$)	Kline and Walters (2016). It is conservative to assume that a 1 standard deviation increase in test scores increases earnings by 10%.
e_U	$291,287	Chetty et al. (2011). The average present discounted value of earnings at age 10 is $522,000 (2010 dollars). Discounted back to age 4 assuming a discount rate of 3% and inflation-adjusted to 2005 dollars.
e_T	$0.8 \times e_U$,	Kline and Walters (2016). Average lifetime earnings of low-income children is lower than the national average.
e_K	$380,064	Chetty et al. (2011). The average present discounted value of earnings at age 10 is $522,000 (2010 dollars). Discounted back to age 5 assuming a discount rate of 3% and inflation-adjusted to 2005 dollars.
τ	0.2	Conservative assumption suggested in personal conversation with Nathan Hendren

Notes: "U" denotes universal pre-K; "T" denotes targeted pre-K, and "K" denotes universal kindergarten. All monetary values are in 2005 dollars unless otherwise noted.

Appendix References

Barnett, W. Steven, Jason T. Hustedt, Laura E. Hawkinson, and Kenneth B. Robin. 2006. *The State of Preschool 2006*. New Brunswick, NJ: The National Institute for Early Education Research.

Barnett, W. Steven, Jason T. Hustedt, Allison H. Friedman, Judi Stevenson Boyd, and Pat Ainsworth. 2007. *The State of Preschool 2007*. New Brunswick, NJ: The National Institute for Early Education Research.

Cascio, Elizabeth. 2017. "Does Universal Preschool Hit the Target? Program Access and Preschool Impacts." IZA Discussion Paper 10596.

Chetty, Raj, John N. 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(4): 1593-1660.

Hendren, Nathaniel. 2016. "The Policy Elasticity." *Tax Policy and the Economy* 30(1): 51-89.

Kline, Patrick and Christopher Walters. 2016. "Evaluating Public Programs with Close Substitutes: The Case of Head Start." *Quarterly Journal of Economics* 131(4): 1795-1848.

Appendix Figure 1. Pre-K Eligibility and Pre-K Attendance by Age, State Program Type, and Poverty Status
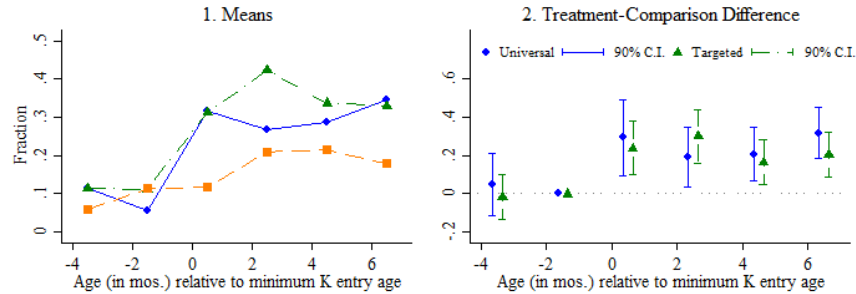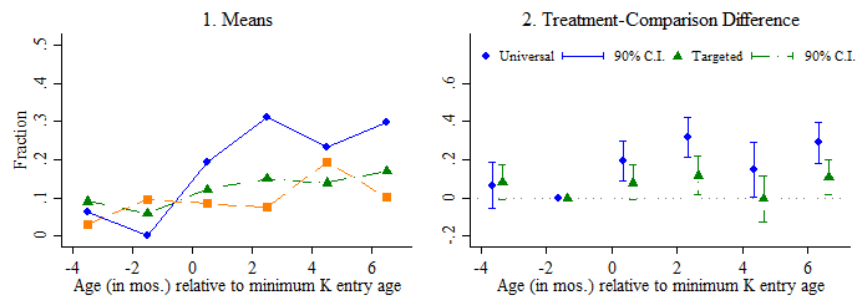
A. Low-Income (<=185% FPL)

B. Not Low-Income (>185% FPL)

Data are from the ECLS-B. Estimation sample is restricted to respondents with non-missing values of key variables resident in one of the analysis states at wave 3 (2005-06), born with 4 months after and 8 months before that state's cutoff birthdate for kindergarten entry, and assessed during the 2005-06 school year. The dependent variable in each panel is a dummy for pre-K attendance during wave 3, when respondents were 4 years of age. Panel A corresponds to respondents who were eligible for free- or reduced-price lunch in 2005-06; Panel B corresponds to respondents who were not. Subpanel 1 of each panel plots the average standardized test score by age relative to the minimum age for kindergarten entry (2-month bins) for treated states with universal pre-K programs, treated states with targeted pre-K programs, and comparison states; see notes to Table 1 or Appendix Tables 1 and 2. The dots in subpanel 2 of each panel represent, separately for treatment states with universal programs and treatment states with targeted programs, the coefficients on interactions between a treatment dummy and a series of dummies for age relative to the minimum age for kindergarten entry (2-month bins) from a regression that allows for direct effects of each of these (sets of) variables in addition to month x year of assessment dummies and state fixed effects. The interaction with the dummy for missing eligibility by 1 to 2 months is omitted for identification. Capped vertical lines represent 90% confidence intervals, with standard errors clustered on state by month of birth.

54

Appendix Figure 2. Pre-K Eligibility and Reading Scores by Age, State Program Type, and Poverty Status

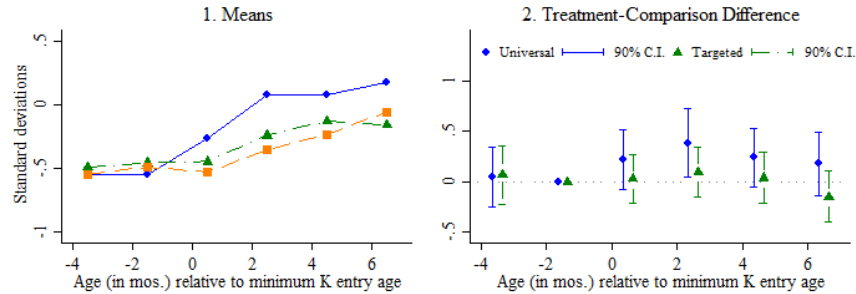### A. Low-Income (<=185% FPL)

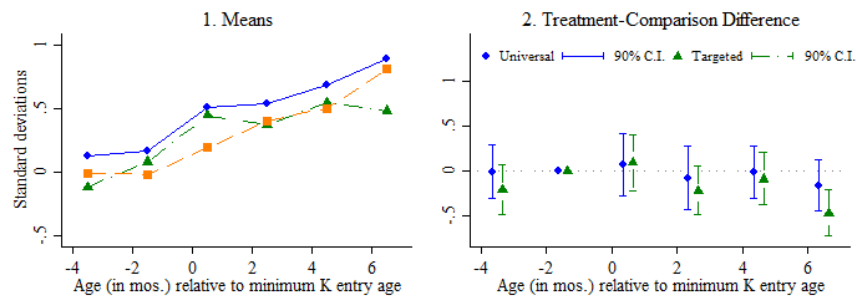### B. Not Low-Income (>185% FPL)

Notes: Data are from the ECLS-B. Estimation sample is restricted to respondents with non-missing values of key variables resident in one of the analysis states at wave 3 (2005-06), born with 4 months after and 8 months before that state's cutoff birthdate for kindergarten entry, and assessed during the 2005-06 school year. The dependent variable in each panel is the standardized reading score during wave 3, when respondents were 4 years of age. Panel A corresponds to respondents who were eligible for free- or reduced-price lunch in 2005-06; Panel B corresponds to respondents who were not. Subpanel 1 of each panel plots the average standardized test score by age relative to the minimum age for kindergarten entry (2-month bins) for treated states with universal pre-K programs, treated states with targeted pre-K programs, and comparison states; see notes to Table 1 or Appendix Tables 1 and 2. The dots in subpanel 2 of each panel represent, separately for treatment states with universal programs and treatment states with targeted programs, the coefficients on interactions between a treatment dummy and a series of dummies for age relative to the minimum age for kindergarten entry (2-month bins) from a regression that allows for direct effects of each of these (sets of) variables in addition to month x year of assessment dummies and state fixed effects. The interaction with the dummy for missing eligibility by 1 to 2 months is omitted for identification. Capped vertical lines represent 90% confidence intervals, with standard errors clustered on state by month of birth.

Appendix Figure 3. Pre-K Eligibility and Math Scores by Age, State Program Type, and Poverty Status

A. Low-Income (<=185% FPL)

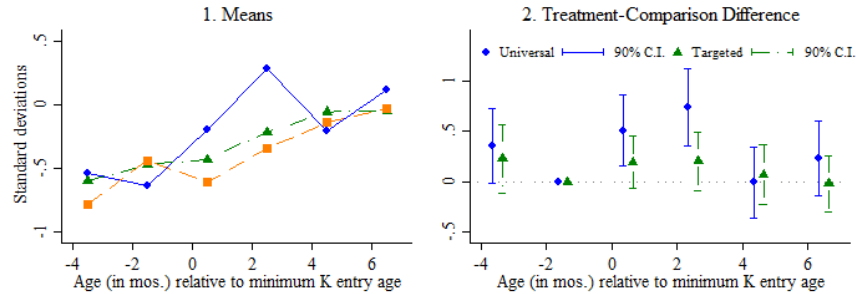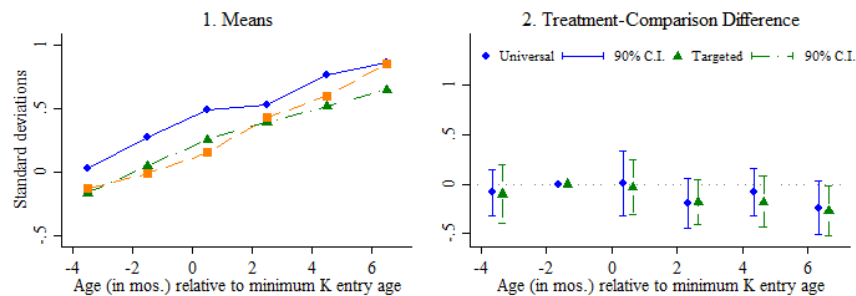### B. Not Low-Income (>185% FPL)

Notes: Data are from the ECLS-B. Estimation sample is restricted to respondents with non-missing values of key variables resident in one of the analysis states at wave 3 (2005-06), born with 4 months after and 8 months before that state's cutoff birthdate for kindergarten entry, and assessed during the 2005-06 school year. The dependent variable in each panel is the standardized math score during wave 3, when respondents were 4 years of age. Panel A corresponds to respondents who were eligible for free- or reduced-price lunch in 2005-06; Panel B corresponds to respondents who were not. Subpanel 1 of each panel plots the average standardized test score by age relative to the minimum age for kindergarten entry (2-month bins) for treated states with universal pre-K programs, treated states with targeted pre-K programs, and comparison states; see notes to Table 1 or Appendix Tables 1 and 2. The dots in subpanel 2 of each panel represent, separately for treatment states with universal programs and treatment states with targeted programs, the coefficients on interactions between a treatment dummy and a series of dummies for age relative to the minimum age for kindergarten entry (2-month bins) from a regression that allows for direct effects of each of these (sets of) variables in addition to month x year of assessment dummies and state fixed effects. The interaction with the dummy for missing eligibility by 1 to 2 months is omitted for identification. Capped vertical lines represent 90% confidence intervals, with standard errors clustered on state by month of birth.

Appendix Figure 4. Kindergarten Attendance and Age 5 Test Scores by Age
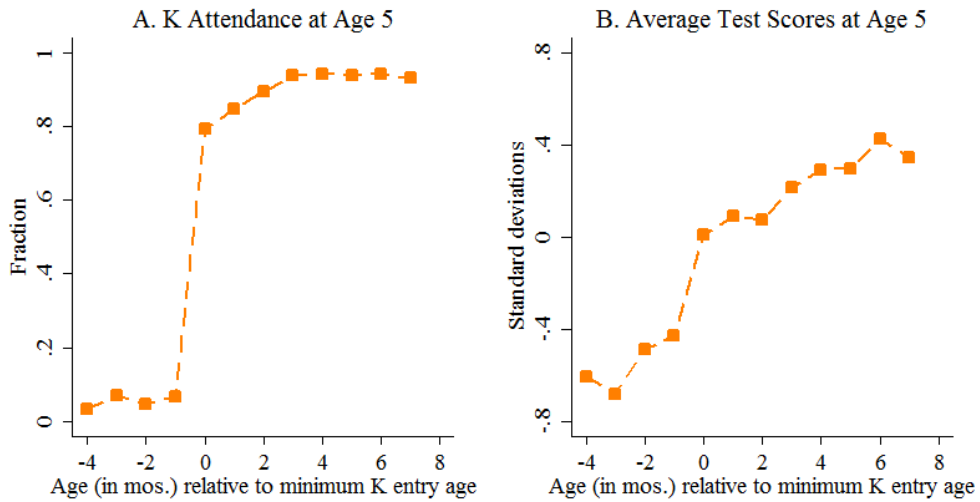
Notes: Data are from the ECLS-B. Estimation sample is restricted the subset of the original estimation sample born within 4 months after and 8 months before birthdate for kindergarten entry in their wave 4 state of residence. The dependent variable in each panel is measured during wave 4, when respondents were 5 years of age. Each panel plots averages by age relative to the minimum age for kindergarten entry across all analysis states (universal pre-K, targeted pre-K, comparison); see notes to Table 1 or Appendix Tables 1 and 2.

Appendix Table 1. Characteristics of State Pre-Kindergarten Programs Under Study

State	Birthday Cutoff for Pre-K	% enrolled by age			Quality Checklist (out of 10)				Annual spending per child		
		age 4	age 3	diff.	Total	Teacher Train. & Creds. (4)	Staffing Ratios & Class Size (2)	Other (4)	Pre-K	Head Start	K-12
					A. Universal Programs						-
Oklahoma	Sept. 1	70.2	0	70.2	9	3	2	4	$6,167	$5,809	$7,475
Georgia	Sept. 1	51.5	0	51.5	8	2	2	4	$3,978	$7,149	$10,492
Florida	Sept. 1	46.5	0	46.5	4	0	2	2	$2,163	$7,386	$9,739
West Virginia	Sept. 1	39.9	4.5	35.4	7	2	2	3	$7,758	$6,637	$11,262
Wisconsin	Sept. 1	32.1	0.7	31.4	5.1	2.9	0.1	2.1	$4,590	$6,695	$12,789
New York	Dec. 1	28.6	0.5	28.1	5.6	1.4	2	2.2	$3,512	$8,794	$15,235
pop-weighted avg.	*-*	*41.5*	*0.3*	*41.2*	*5.8*	*1.3*	*1.8*	*2.6*	*$3,569*	*$7,627*	*$11,875*
					B. Targeted Programs						-
Texas	Sept. 1	44.3	4.5	39.8	4	3	0	1	$2,653	$7,091	$9,076
South Carolina	Sept. 1	31	4.2	26.8	8	3	2	3	$3,219	$6,718	$10,542
Maryland	Sept. 1	30.7	1	29.7	7	3	2	2	$4,663	$7,522	$10,773
Illinois	Sept. 1	23	14.4	8.6	9	4	2	3	$3,298	$6,812	$11,402
Louisiana	Sept. 30	21.6	0	21.6	7.8	2.05	2	3.75	$5,012	$6,620	$10,241
Michigan	Dec. 1	16.2	0	16.2	6	3	2	1	$3,934	$6,670	$11,437
Kansas	Aug. 31	14.5	0	14.5	3	2	0	1	$2,554	$6,404	$9,676
Colorado	Oct. 1	13.5	2.2	11.3	4	1	2	1	$3,056	$6,941	$9,877
Virginia	Sept. 30	11.1	0	11.1	7	2	2	3	$5,375	$7,216	$11,626
Tennessee	Sept. 30	10.6	0.5	10.1	9	3	2	4	$4,061	$7,238	$7,561
pop-weighted avg.	*-*	*27.2*	*4.1*	*23.1*	*6.2*	*2.9*	*1.3*	*2.0*	*$3,537*	*$6,974*	*$10,139*

Notes: Unless otherwise noted, source is Barnett, et al. (2006), and figures correspond to the 2005-06 academic year. Monetary figures are in nominal dollars. The other components of the quality checklist come from comprehensive early learning standards, comprehensive services provided (vision, hearing, health, and one support service, at least one meal), and a site visit requirement; see notes to Figure 2 for complete description of the checklist. Head Start spending corresponds to 2004-05, and K-12 spending includes both current and capital expenditures. Spending on pre-K is a lower bound, since programs may receive funding from local or federal sources not reported.

Appendix Table 2. Characteristics of the Comparison States

State	Birthday Cutoff for Pre-K and K	% enrolled in state funded pre-K by age		
		age 4	age 3	diff.
		A. Comparison States with Pre-K Programs		
Alabama	Sept. 1	1.7	0	1.7
California	Dec. 2	9.9	4.5	5.4
Delaware	Aug. 31	7.8	0	7.8
Missouri	Jul. 31	4	2.3	1.7
New Mexico	Sept. 1	6.8	0.6	6.2
Ohio	Sept. 30	4.4	1	3.4
Oregon	Sept. 1	5	2.6	2.4
Washington	Aug. 31	6	1.4	4.6
pop-weighted avg.	-	*7.5*	*3.0*	*4.4*
		B. Comparison States without Pre-K Programs		
Alaska	Sept. 1	0	0	0
Hawaii	Dec. 31	0	0	0
Idaho	Sept. 1	0	0	0
Indiana	Aug. 1	0	0	0
Mississippi	Sept. 1	0	0	0
North Dakota	Aug. 31	0	0	0
Rhode Island	Sept. 1	0	0	0
South Dakota	Sept. 1	0	0	0
Utah	Sept. 1	0	0	0
pop-weighted avg.	-	*0.0*	*0.0*	*0.0*
		C. All Comparison States		
pop-weighted avg.	-	*5.9*	*2.4*	*3.5*

Notes: Unless otherwise noted, source is Barnett, et al. (2006, 2007).

Appendix Table 3. Descriptive Statistics and Balance Tests on Key Variables, by Family Income and Program Type

	Children:	Low-income				Not Low-Income			
	States:	Universal		Targeted		Universal		Targeted	
		Mean	Coef. (se)	Mean	Coef. (se)	Mean	Coef. (se)	Mean	Coef. (se)
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A. Treatment variable									
Pre-kindergarten[a]		0.081	0.227	0.113	0.226	0.025	0.206	0.074	0.041
			(0.063)		(0.050)		(0.050)		(0.038)
B. Background characteristics									
Age in months[a]		48.551	-0.042	48.516	-0.107	48.502	-0.146	48.003	-0.077
		[3.060]	(0.074)	[3.224]	(0.072)	[3.196]	(0.076)	[3.359]	(0.060)
Female		0.493	0.039	0.512	0.009	0.436	0.086	0.529	-0.011
			(0.082)		(0.064)		(0.057)		(0.055)
Black non-Hispanic		0.349	0.014	0.345	0.031	0.077	0.065	0.082	0.042
			(0.060)		(0.055)		(0.029)		(0.026)
Hispanic		0.206	0.000	0.375	-0.066	0.188	-0.038	0.156	0.037
			(0.059)		(0.050)		(0.047)		(0.045)
Low birth weight		0.075	0.043	0.096	0.030	0.073	0.014	0.065	0.012
			(0.025)		(0.028)		(0.020)		(0.019)
Maternal education ≤ HS[a]		0.663	0.069	0.695	0.035	0.249	0.054	0.281	-0.025
			(0.061)		(0.068)		(0.063)		(0.051)
Both biological parents in HH [a]		0.511	-0.063	0.547	0.003	0.877	-0.113	0.812	0.020
			(0.057)		(0.056)		(0.052)		(0.051)
Non-English at home[a]		0.199	-0.002	0.243	-0.005	0.111	0.013	0.091	0.014
			(0.050)		(0.041)		(0.036)		(0.041)
p-value: joint test for background chars			0.61		0.44		0.04		0.41
Observations[b]		150	1550	200	1750	150	1850	250	2200

Notes: A child is considered low income if his (preschool-age or 2005-06) family income is at or below 185% FPL, the threshold for eligibility for reduced-price lunch and the modal income-eligibility criterion for the targeted programs under study. Odd-numbered columns give means for respondents in treatment states ineligible for pre-K in 2005-06 [standard deviations for non-binary variables]. Even-numbered columns give DD coefficients (standard errors) on the interaction between a dummy for being eligible for pre-K in 2005-06 and a dummy for residing in a treatment state (*elig* x *treat*) from a separate regression that also includes dummies for state of residence, month x year of assessment, and month age five relative to the state kindergarten entry cutoff birthdate in 2006-07. Treatment states are those with state-funded pre-K programs focused much more on 4-year-olds than 3-year-olds and statewide minimum age at pre-K entry cutoffs not in the middle of the month; treatment states with universal programs are FL, GA, NY, OK, WI, and WV, and treatment states with targeted programs are CO, IL, KS, LA, MI, MD, SC, TN, TX, and VA. Comparison states have statewide age at kindergarten entry regulations; some comparison states have relatively small pre-K programs (AL, CA, DE, MO, NM, OH, OR, WA), while others lack pre-K programs (AK, HI, ID, IN, MS, ND, RI, SD, UT). A child is deemed eligible for K in 2006-07 if he /she turned age 5 in time to start K in fall 2006, given his/her date of birth and the kindergarten entry age regulations in effect in 2006-07 reported by Barnett et al. (2007). Sample is limited to children who turn age 5 between 4 months after and 8 months before the cutoff date and who are assessed during the school year. Means and regressions are weighted by longitudinal sampling weights, and standard errors (in parentheses) are clustered on state x month of birth. [a] Measured at preschool age, or in 2005-06 (wave 3 interview). [b] rounded to the nearest 50, per IES guidelines.

Appendix Table 4. Sensitivity of Estimated Effects of Pre-K to Choice of Outcome: Low-Income Subsample

	Pre-K: First Stage (1)	Test Scores RF (ITT) (2)	IV (TOT) (3)	OLS (4)
A. Baseline				
Universal (N=1,550)	0.226	0.263	1.160	-0.037
	(0.062)	(0.107)	(0.544)	(0.055)
Targeted (N=1,750)	0.223	-0.018	-0.082	-0.015
	(0.051)	(0.102)	(0.450)	(0.051)
p-value on difference	0.966	0.020	0.032	0.728
B. Reading score only				
Universal (N=1,550)	0.226	0.272	1.201	0.028
	(0.062)	(0.109)	(0.551)	(0.056)
Targeted (N=1,750)	0.223	-0.039	-0.177	-0.007
	(0.051)	(0.102)	(0.447)	(0.058)
p-value on difference	0.966	0.012	0.021	0.593
C. Math score only				
Universal (N=1,550)	0.226	0.253	1.118	-0.102
	(0.062)	(0.130)	(0.629)	(0.068)
Targeted (N=1,750)	0.223	0.003	0.013	-0.024
	(0.051)	(0.117)	(0.515)	(0.056)
p-value on difference	0.966	0.080	0.096	0.286
D. Mental score (age 2)[a]				
Universal (N=1,550)	0.234	-0.006	-0.025	-0.038
	(0.064)	(0.089)	(0.373)	(0.060)
Targeted (N=1,750)	0.207	0.159	0.772	-0.117
	(0.051)	(0.103)	(0.519)	(0.062)
p-value on difference	0.701	0.063	0.068	0.111
E. = 1 if parent reports not ready for K (age 4)				
Universal (N=1,550)	0.229	-0.143	-0.623	0.002
	(0.062)	(0.073)	(0.357)	(0.029)
Targeted (N=1,750)	0.223	-0.038	-0.172	-0.081
	(0.051)	(0.057)	(0.256)	(0.026)
p-value on difference	0.932	0.188	0.220	0.010

Notes: A child is considered low income if his (preschool-age or 2005-06) family income is at or below 185% FPL, the threshold for eligibility for reduced-price lunch and the modal income-eligibility criterion for the targeted programs under study. The first stage and RF (ITT) (reduced form (intent-to-treat)) columns give coefficients on the interaction between a dummy for being eligible for kindergarten in 2006-07 (same as a dummy for being eligible for pre-K in 2005-06 in a treatment state) and a dummy for being in a treated state (*elig* x *treat*). (See Table 1 or Data Appendix for a definition of treatment and comparison states.) The IV (TOT) (instrumental variables (treatment on treated)) and OLS columns give coefficients on pre-K attendance (*prek*), in the first case instrumenting with *elig* x *treat* using TSLS. All coefficients are from separate regressions that also include as controls dummies for state of residence, month x year of assessment, and month age five relative to the state kindergarten entry cutoff birthdate in 2006-07, and all demographic and background characteristics listed in Table 1 Panel B except the low-income indicator. Regressions are weighted by longitudinal sampling weights, and standard errors (in parentheses) are clustered on state x month of birth. [a] Additional controls include age at assessment and dummies for month x year of assessment in wave 2 of the ECLS-B, and panel weights incorporate observation in wave 2.

Appendix Table 5. Sensitivity of Estimated Effects of Pre-K on Test Scores to Controlling for Age 2 Scores

	Pre-K: First Stage (1)	Test Scores		
		RF (ITT) (2)	IV (TOT) (3)	OLS (4)
		A. Full Sample		
Universal (N=3,400)	0.210	0.128	0.609	-0.054
	(0.046)	(0.063)	(0.313)	(0.045)
Targeted (N=3,950)	0.112	-0.009	-0.084	-0.067
	(0.034)	(0.065)	(0.571)	(0.040)
p-value on difference	0.043	0.038	0.178	0.790
		B. Low-Income Subsample		
Universal (N=1,550)	0.232	0.252	1.086	-0.016
	(0.063)	(0.102)	(0.505)	(0.053)
Targeted (N=1,750)	0.206	0.002	0.008	0.012
	(0.050)	(0.104)	(0.495)	(0.050)
p-value on difference	0.707	0.028	0.055	0.653

Notes: A child is considered low income if his (preschool-age or 2005-06) family income is at or below 185% FPL, the threshold for eligibility for reduced-price lunch and the modal income-eligibility criterion for the targeted programs under study. The first stage and RF (ITT) (reduced form (intent-to-treat)) columns give coefficients on the interaction between a dummy for being eligible for kindergarten in 2006-07 (same as a dummy for being eligible for pre-K in 2005-06 in a treatment state) and a dummy for being in a treated state (*elig* x *treat*). (See Table 1 or Data Appendix for a definition of treatment and comparison states.) The IV (TOT) (instrumental variables (treatment on treated)) and OLS columns give coefficients on pre-K attendance (*prek*), in the first case instrumenting with *elig* x *treat* using TSLS. All coefficients are from separate regressions that also include as controls dummies for state of residence, month x year of assessment, and month age five relative to the state kindergarten entry cutoff birthdate in 2006-07, all demographic and background characteristics listed in Table 1 Panel B except the low-income indicator, age 2 motor and mental scores, age at the wave 2 assessment, and dummies for month x year of wave 2 assessment. Missing values of wave 2 assessments are imputed with (weighted) sample means and indicators for imputation also included as regressors. Regressions are weighted by longitudinal sampling weights, and standard errors (in parentheses) are clustered on state x month of birth.

Appendix Table 6. IV Estimates of the Impact of Pre-K Attendance on Parental Perceptions of Kindergarten Readiness, by Reason: Low-Income Subsample

	Reason Not Ready for Kindergarten (=0 if deemed ready):				
	Academic	Social	Behavioral	Physical	Speech Concerns
	(1)	(2)	(3)	(4)	(5)
Universal (N=1,550)	-0.265	0.010	-0.003	0.046	-0.158
	(0.170)	(0.093)	(0.145)	(0.044)	(0.094)
Targeted (N=1,750)	-0.044	0.104	-0.080	0.057	-0.073
	(0.077)	(0.086)	(0.130)	(0.047)	(0.097)
p-value on difference	0.173	0.268	0.598	0.616	0.442

	Medical	Limited English	Young in Class	Limited Preschool	
	(6)	(7)	(8)	(9)	
Universal (N=1,550)	-0.063	0.016	-0.038	-0.050	
	(0.077)	(0.030)	(0.031)	(0.027)	
Targeted (N=1,750)	0.001	-0.011	0.002	-0.039	
	(0.051)	(0.019)	(0.006)	(0.023)	
p-value on difference	0.394	0.309	0.213	0.569	

Notes: Dependent variables are coded as zero for children of parents who report no concern over their kindergarten readiness. Not every parent who reports concern over kindergarten readiness gives a reason – and parents can give multiple reasons – so the coefficients need not add up to those reported in Appendix Table 4 Panel E. Each cell entry represents an IV estimate of the impact of pre-K attendance on a dummy for a parent reporting (in wave 3) concerns that a child is not ready for kindergarten for the reason given. The instrument for pre-K attendance is the interaction between a dummy for being eligible for kindergarten in 2006-07 (same as a dummy for being eligible for pre-K in 2005-06 in a treatment state) and a dummy for being in a treated state (elig x treat). All underlying regressions include dummies for for state of residence, month x year of assessment, and month age five relative to the state kindergarten entry cutoff birthdate in 2006-07, and all demographic and background characteristics listed in Table 1 Panel B. Regressions are weighted by longitudinal sampling weights, and standard errors (in parentheses) are clustered on state x month of birth.

Appendix Table 7. Balance Tests on Key Variables, by Program Type:
+/- 4 Months from Cutoff

	Universal	Targeted	Uni - Tar
	DD Coef.	DD Coef.	DDD Coef.
	(se)	(se)	(se)
	(1)	(2)	(3)
A. Treatment variable			
Pre-kindergarten[a]	0.221	0.125	0.096
	(0.055)	(0.038)	(0.056)
B. Background characteristics			
Age in months[a]	-0.156	-0.140	-0.016
	(0.063)	(0.054)	(0.057)
Female	0.117	0.060	0.057
	(0.053)	(0.050)	(0.057)
Black non-Hispanic	0.030	0.016	0.015
	(0.041)	(0.032)	(0.042)
Hispanic	-0.021	-0.011	-0.010
	(0.046)	(0.037)	(0.042)
Low birth weight	0.028	0.012	0.015
	(0.016)	(0.013)	(0.016)
Maternal education ≤ HS[a]	0.012	-0.048	0.060
	(0.044)	(0.050)	(0.044)
Both biological parents in HH [a]	-0.079	0.025	-0.105
	(0.047)	(0.041)	(0.051)
Non-English at home[a]	-0.009	-0.024	0.016
	(0.034)	(0.032)	(0.036)
Family income ≤ 185% FPL[a]	-0.075	-0.080	0.005
	(0.047)	(0.048)	(0.051)
C. _p_-value: joint test for background chars.			
All background characteristics	0.01	0.17	0.26
Excluding poverty	0.32	0.40	0.27
Observations[b]	2,000	2,350	4,350

Notes: Sample is limited to children who turn age 5 between 4 months after and 4 months before the cutoff date and who are assessed during the school year. Each column gives DD coefficients (standard errors) on the interaction between a dummy for being eligible for pre-K in 2005-06 and a dummy for residing in a treatment state (_elig_ x _treat_) from a separate regression that also includes dummies for state of residence, month x year of assessment, and month age five relative to the state kindergarten entry cutoff birthdate in 2006-07. Treatment states are those with state-funded pre-K programs focused much more on 4-year-olds than 3-year-olds and statewide minimum age at pre-K entry cutoffs not in the middle of the month; treatment states with universal programs are FL, GA, NY, OK, WI, and WV, and treatment states with targeted programs are CO, IL, KS, LA, MI, MD, SC, TN, TX, and VA. Comparison states have statewide age at kindergarten entry regulations; some comparison states have relatively small pre-K programs (AL, CA, DE, MO, NM, OH, OR, WA), while others lack pre-K programs (AK, HI, ID, IN, MS, ND, RI, SD, UT). A child is deemed eligible for K in 2006-07 if he /she turned age 5 in time to start K in fall 2006, given his/her date of birth and the kindergarten entry age regulations in effect in 2006-07 reported by Barnett et al. (2007). Means and regressions are weighted by longitudinal sampling weights, and standard errors (in parentheses) are clustered on state x month of birth. [a] Measured at preschool age, or in 2005-06 (wave 3 interview). [b] rounded to the nearest 50, per IES guidelines.

Appendix Table 8. Sensitivity of Estimated Effects of Pre-K on Test Scores to Estimation Sample: Low-Income Subsample

	Pre-K: First Stage	Test Scores		
		RF (ITT)	IV (TOT)	OLS
	(1)	(2)	(3)	(4)
A. Baseline				
Universal (N=1,550)	0.226	0.263	1.160	-0.037
	(0.062)	(0.107)	(0.544)	(0.055)
Targeted (N=1,750)	0.223	-0.018	-0.082	-0.015
	(0.051)	(0.102)	(0.450)	(0.051)
p-value on difference	0.966	0.020	0.032	0.728
B. +/- 4 months from threshold[a]				
Universal (N=900)	0.242	0.347	1.434	-0.042
	(0.083)	(0.129)	(0.677)	(0.078)
Targeted (N=1,000)	0.241	0.049	0.202	0.026
	(0.061)	(0.112)	(0.443)	(0.075)
p-value on difference	0.985	0.029	0.060	0.474
C. Expanded sample (+ middle of month cutoffs)[b]				
Universal (N=1,600)	0.230	0.249	1.080	-0.023
	(0.061)	(0.107)	(0.527)	(0.055)
Targeted (N=1,900)	0.211	0.008	0.036	0.015
	(0.048)	(0.095)	(0.443)	(0.049)
p-value on difference	0.776	0.037	0.059	0.521
D. Restricted sample (- late cutoffs)[c]				
Universal (N=950)	0.302	0.182	0.601	0.002
	(0.066)	(0.116)	(0.376)	(0.098)
Targeted (N=1,150)	0.303	-0.050	-0.164	0.001
	(0.049)	(0.116)	(0.377)	(0.077)
p-value on difference	0.996	0.037	0.033	0.994

Notes: A child is considered low income if his (preschool-age or 2005-06) family income is at or below 185% FPL, the threshold for eligibility for reduced-price lunch and the modal income-eligibility criterion for the targeted programs under study. The first stage and RF (ITT) (reduced form (intent-to-treat)) columns give coefficients on the interaction between a dummy for being eligible for kindergarten in 2006-07 (same as a dummy for being eligible for pre-K in 2005-06 in a treatment state) and a dummy for being in a treated state (*elig* x *treat*). (See Table 1 or Data Appendix for a definition of treatment and comparison states in the baseline sample.) The IV (TOT) (instrumental variables (treatment on treated)) and OLS columns give coefficients on pre-K attendance (*prek*), in the first case instrumenting with *elig* x *treat* using TSLS. All coefficients are from separate regressions that also include as controls dummies for state of residence, month x year of assessment, and month age five relative to the state kindergarten entry cutoff birthdate in 2006-07, and all demographic and background characteristics listed in Table 1 Panel B except the low-income indicator. Regressions are weighted by longitudinal sampling weights, and standard errors (in parentheses) are clustered on state x month of birth. [a] Sample further limited to respondents with birthdays within 4 months of the cutoff birthdate for kindergarten entry in their wave 3 state of residence. [b] Sample expanded to include respondents residing in treatment states with middle-of-month cutoffs (targeted: AR, NC; universal: ME) not born in the cutoff birthdate month. [c] Sample eliminates treatment and comparison states with cutoff birthdates after October 1 (universal: NY, targeted: MI, comparison: CA, HI).

Appendix Table 9. Impacts on Alternative Care Arrangements, by State Program Type: Full Sample

	Pre-K: First Stage (1)	Head Start (2)	Other center-based care (3)	Any formal (2+3) (4)	Informal non-parental care (5)	Parental care (6)	Any informal (5+6) (7)
				Alternatives:			
				A. Ineligible Means			
Universal	0.05	0.12	0.44	0.56	0.14	0.25	0.39
Targeted	0.09	0.12	0.34	0.46	0.18	0.27	0.45
				B. Reduced form (coef (se) on *elig* x *treat$_s$*)			
Universal (N=3,400)	0.211	-0.019	-0.095	-0.114	-0.002	-0.095	-0.097
	(0.047)	(0.035)	(0.055)	(0.066)	(0.029)	(0.040)	(0.045)
Targeted (N=3,950)	0.118	-0.013	0.001	-0.011	-0.029	-0.078	-0.107
	(0.034)	(0.031)	(0.041)	(0.046)	(0.034)	(0.039)	(0.040)
p-value on difference	0.057	0.871	0.070	0.122	0.436	0.700	0.851
				C. Instrumental Variables (coef (se) on *prek$_{is}$*)			
Universal (N=3,400)	n.a.	-0.090	-0.448	-0.539	-0.012	-0.450	-0.461
		(0.155)	(0.219)	(0.239)	(0.136)	(0.219)	(0.239)
Targeted (N=3,950)	n.a.	-0.107	0.011	-0.096	-0.245	-0.659	-0.904
		(0.248)	(0.343)	(0.371)	(0.294)	(0.332)	(0.371)
p-value on difference	n.a.	0.947	0.159	0.222	0.385	0.511	0.222

Notes: Reduced-form coefficients in Panel A are on the interaction between a dummy for being eligible for kindergarten in 2006-07 (same as a dummy for being eligible for pre-K in 2005-06 in a treatment state) and a dummy for being in a treated state (*elig* x *treat*). (See Table 1 or Data Appendix for a definition of treatment and comparison states.) The instrumental variables coefficients in Panel B are on pre-K attendance (*prek*), estimated by instrumenting with *elig* x *treat* using TSLS. All coefficients are from separate regressions that also include as controls dummies for state of residence, month x year of assessment, and month age five relative to the state kindergarten entry cutoff birthdate in 2006-07, and all demographic and background characteristics listed in Table 1 Panel B. Regressions are weighted by longitudinal sampling weights, and standard errors (in parentheses) are clustered on state x month of birth. "n.a." means "not applicable."

Appendix Table 10. Impacts on Alternative Care Arrangements to Kindergarten: Full Sample

	Public K: First Stage (1)	Private K (2)	Grades 1-2 (3)	Head Start (4)	Alternatives: Other center-based care (5)	Any formal (4+5) (6)	Informal non-parental care (7)	Parental care (8)	Any informal (7+8) (9)
					A. Ineligible means				
All states	0.05	0.02	0.00	0.13	0.58	0.71	0.11	0.11	0.21
					B. Reduced form (coef (se) on *elig*)				
All states (N=2,400)	0.657 (0.036)	0.070 (0.024)	0.000 (0.000)	-0.158 (0.033)	-0.319 (0.048)	-0.477 (0.045)	-0.179 (0.037)	-0.071 (0.031)	-0.250 (0.038)
					C. Instrumental Variables (coef (se) on kindergarten attendance)				
All states (N=2,400)	n.a.	0.107 (0.040)	0.000 (0.000)	-0.241 (0.049)	-0.485 (0.068)	-0.726 (0.060)	-0.273 (0.058)	-0.108 (0.046)	-0.381 (0.059)

Notes: Reduced-form coefficients in Panel A are on a dummy for being eligible for kindergarten in 2006-07 (*elig*). The instrumental variables coefficients in Panel B are give coefficients on K attendance, instrumenting with *elig* using TSLS. All coefficients are from separate regressions that also include as controls dummies for state of (wave 4) residence and month x year of (wave 4) assessment, a linear term in month age five relative to the state kindergarten entry cutoff birthdate in 2006-07, an interaction between that linear term and *elig*, and all demographic and background characteristics listed in Table 1 Panel B. Regressions are weighted by longitudinal sampling weights, and standard errors (in parentheses) are clustered on state x month of birth. Sample is limited to the subset of the original estimation sample born within 4 months after and 4 months before the minimum kindergarten entry age in their wave 4 state of residence. "n.a." means "not applicable."